EVALUATING A COMPUTERIZED AID FOR CONDUCTING A COGNITIVE TASK ANALYSIS

by

JEFFREY R. VOIGT
B.S. United States Military Academy, 1987

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science
in the Department of Industrial Engineering and Management Systems
in the College of Engineering
at the University of Central Florida
Orlando, FL

Spring Term
2000

ABSTRACT

Models that simulate cognitive processes have demonstrated considerable success in a variety of technical domains such as, intelligent tutoring, predicting the complexity of human system interactions, decision support and expert systems among others. Creating such models requires considerable skill in conducting a cognitive task analysis. The conduct of a cognitive task analysis is costly and labor intensive. As a result, a few computerized aids have been developed to assist in the process of conducting such analyses. However, none have been evaluated to determine how accurately and consistently users of such tools can create cognitive models. If such tools cannot demonstrate the creation of accurate models of cognitive tasks across users then such tools will be ineffective and unreliable. The research reported herein presents the results of experimentation, which focuses upon the evaluation of a computerized aid, specifically CAT-HCI (Cognitive Analysis Tool – Human Computer Interface), for the conduct of a detailed cognitive task analysis. A sample of users for a newly developed interface (tactical display) for the Army's Bradley A3 Fighting Vehicle were asked to model their knowledge of a routine task. Measures of the accuracy and of the consistency of the user generated models were recorded and analyzed. Accuracy measures the level of agreement between subject models and a baseline model, while consistency measures the level of agreement between subject models.

The experiment found that domain experts, with little understanding of knowledge engineering techniques, used CAT-HCI to generate cognitive models that were on average 76.88% accurate when compared to the baseline. Additionally, the models were broken down into the various subsystems of cognitive, perceptual, and physical for the purpose of further evaluation. This accuracy of between 75%-80% represents the potential savings of three-quarters of a cognitive analysts time, resulting in less costly development for intelligent tutoring systems for training human system interactions.

ACKNOWLEDGMENTS

TRADOC System Manager Bradley Fighting Vehicle System (TSM-Bradley), United States Infantry Center (USAIC), Fort Benning, GA, funded the research contained in this thesis. They supported travel efforts and provided much needed support in coordinating this effort. Additionally, Mr. Donald Sayers, from Director of Training, USAIC, Fort Benning, GA proved invaluable in coordinating the subject matter experts for use in this experiment. I express my sincere gratitude to TSM-Bradley for sponsoring this research and Mr. Don Sayers for his considerable effort in providing subject matter experts.

United Defense Limited Partners (UDLP), Orlando, provided the necessary simulators necessary for data collection at Fort Hood and here in Orlando. Ms. Angela Alban provided much insight into the Bradley Desktop Trainer, and was very helpful in coordinating the workstations and ensuring they were available and functioning when needed. I thank UDLP and Ms. Angela Alban for their support and for providing the workstation for data collection.

I thank Dr. Linda Malone and Dr. Eduardo Sales for their guidance, feedback, and support throughout my research efforts. I thank Dr. Kent Williams for his insight, patience and guidance throughout the research process. His expertise and wisdom have been invaluable during each step of this effort.

I am also deeply indebted to my wonderful wife and three children. Ann's loving support and patience during these two years have been immeasurable. Without her, none of this would have been possible. My three children, Emily, Elizabeth, and Jack, are each a precious gift from God. Their warm hugs, happy smiles, and unconditional love have filled this journey with joy and many great memories.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

BC. Bradley Commander

BDT. (Bradley Desktop Trainer) Bradley A3 Simulation trainer for digital tasks.

BFV Bradley Fighting Vehicle

BFVS Bradley Fighting Vehicle System

Bradley Infantry Fighting Vehicle (BIFV). Provides mobile protected transport of sufficient Infantry to the critical point on the battlefield, fires to support dismounted Infantry, fires to suppress or destroy enemy IFV's and light armor vehicles, and Anti-armor fires to destroy enemy armor. (FM 7-7J, May 93)

Bradley A3. The fourth generation Bradley, scheduled to begin fielding at the beginning of 2000. This fully digitized BIFV will significantly improve the Infantry and Calvary's lethality and command and control on the battlefield.

CAT. Cognitive Analysis Tool. Developed and tested by Kotnour in 1992. A cognitive analysis tool designed to allow domain experts to build cognitive models with minimal need for knowledge engineers.

CAT-HCI. Cognitive Analysis Tool-Human Computer Interaction. Designed by Williams based on lessons learned from CAT.

CHS. Commander's Handstation

CSCP. Commander's Sight Control Panel

CTD. Commander's Tactical Display. A 10.5" flat panel display used by the Bradley commander to interact with the Bradley's digital capability.

CTED. Commander's Tactical Entry Device (Keyboard)

GOMS. Goals, Operators, Methods, and Decision Rules. A methodology for designing cognitive models, well suited for procedural tasks.

Information Dominance. The Army's strategy of using modern technology and information resources to provide our warfighters a decisive edge on the battlefield through improved command and control and situational awareness.

RBD. Remote Biocular Display

SCB. System Control Box

Situational Awareness. Knowing where you are, where other friendly elements are and where the enemy is. A key component of Information Dominance.

SME. Subject Matter Expert (Used synonymously with Domain Expert)

UDLP. United Defense Limited Partners – Prime Contractor for the Bradley A3

1 INTRODUCTION

1.1 Information Dominance

As we enter the 21st century, the United States Military is continuing to modernize equipment and develop strategies that will ensure its supremacy as the world's remaining superpower. In the past, much of the effort for maintaining dominance over the enemy was based on superior firepower through improved munitions. With currently constrained military budgets, especially in the area of modernization, the Army has chosen Information Dominance as the cornerstone of its Army XXI strategy that will carry it into the new millennium and ever-changing strategic environment. According to the Army Digitization Office, Tenants of Army XXI include:

1. Flexible Engagement Strategy
2. 21st Century Technology
3. Knowledge and Capabilities Based
4. Split Based Operations
5. Improved Lethality, Survivability and Tempo
6. Shared Situational Awareness
7. Real-time Information

These tenants combine to form a strategy that will provide our warfighters a decisive edge on the battlefield through Information Dominance. The key enabler to achieving Information Dominance is digitization. The Army Digitization Office explains; "digitizing the Battlefield applies information technologies to acquire, exchange, and employ timely digital information throughout the battlespace, tailored to the needs of each decider (commander), shooter, and supporter...allowing each to maintain a clear and accurate vision of his battlespace necessary to support both planning and execution." Digitization will have a far-reaching impact by improving 1) command and control, 2) situational awareness, 3) common relevant picture, and 4) logistics management.

Command and Control (C2) - A seamless vertical and horizontal information flow will provide more timely dissemination of information up and down the chain of command as well as across the battlefield. Additionally, commanders will have enhanced C2 planning and execution tools that should improve the decision making process. Finally, digital communications provide a much more efficient method of communication.

Situational Awareness (SA) – In its simplest form, SA answers these three questions: 1) Where am I? 2) Where are my buddies? 3) Where is the enemy? This information is highly relevant to every soldier from the rifleman to the Brigade Commander. In the past, 80% of voice communications dealt with these three questions. The Global Positioning System (GPS), which helps answer the first question, has been attributed with much of the success during Operation Desert Storm by providing an accurate location of our soldiers. The Iraqi military did not expect the coalition forces to be able to navigate in the desert sands with so few landmarks. The GPS was relatively new and generally found only at company level and above. It was the beginning of answering the first question, "where am I?" By knowing where they were, units were able to locate and seize their objectives, which ultimately lead to victory.

Currently, GPS and other navigation systems are found at individual vehicle/platform level. Through digitization, vehicles are able to transmit their own vehicle location to other friendly units in the area, thereby fully answering the first two questions. The third question, "where is the enemy?" is much more difficult to answer, but thanks to laser target designators, which can pinpoint identified enemy locations, and near real-time digital communications, observed enemy targets can be rapidly distributed across the force.

Common Relevant Picture – A common relevant picture will provide both a better view of the current and future battlefield and better synchronization of the battle plans.

Logistical Management – Digitization will provide improved asset visibility and assist in cross leveling of assets.

The Army will achieve Information Dominance by providing SA to soldiers at all levels through tactical computers connected by a Tactical Internet. At the tactical level, FBCB2/EBC (Force XXI Battle Command, Brigade and Below/Embedded Battle Command) will provide a more efficient means of data transfer and a better method of aggregation for use by higher-level staffs. FBCB2 provides an interface, set of tools, and message formats that comply with DOD (Department of Defense) standards, yet are pertinent to the leaders and soldiers at Brigade level and below. Some reports, such as a vehicle's location, are automatically generated and forwarded both laterally and vertically to provide a precise picture of how the force is distributed. Each platform's tactical display will graphically portray the terrain superimposed with the updated position reports displayed as icons in near real-time. This not only allows the commander to have a better visual picture of his unit array, it also facilitates the exchange of adjacent unit vehicle locations, resulting in reduced fratricide. Other reports are merely computer assisted. For example, to report an enemy position, the soldier inputs the non-computer-generated data into a preformatted message for submission to appropriate leaders and peers. At the next level this information can be easily reviewed, stored and if necessary, forwarded with the mere push of a button to the next level where data from various sources can be aggregated and analyzed for use by decision-makers.

1.2 The Bradley A3

The Bradley Fighting Vehicle, as introduced into the US Army in 1983, provides critical mobility, protection and firepower to the infantry. It has become a critical member of the combined arms team, allowing the Infantry and supporting tanks to dominate the modern battlefield as evidenced in the 1991 Gulf War. It was during the Gulf War that the United States Army Infantry Center was able for the first time to evaluate the platform's performance under combat conditions. While the US forces enjoyed a resounding victory and the Bradley exceeded expectations, some deficiencies were noted:

1. Navigational problems
2. Inadequate Fire Control System
3. Insufficient space for dismounted soldiers
4. Limited visibility driving
5. Fratricide
6. Not digitized in accordance with the developing weapon systems
7. Could not engage at maximum range of weapon systems
8. Lacked Situational Awareness (SA)
9. No ballistic solution
10. No hunter/killer capability (as found on the M1A2 Abrams tank which allows the commander to search for targets while the gunner is engaging and subsequently hand them off to the gunner)

Recognizing these deficiencies, the Army set out to improve the Infantry Fighting Vehicle. As a short-term solution, several deficiencies were corrected with quickly applied "off the shelf technologies" to produce the Operation Desert Storm (ODS) Bradley which began fielding in 1997. In order to address *all* the above noted deficiencies, the Bradley A3 is being produced to provide a significant technological step forward. Initial fielding is scheduled to begin in 2000. In addition to addressing the deficiencies of the Gulf War, the Bradley A3 will be a completely modernized vehicle and will support the Army's thrust towards digitization in order to achieve Information Dominance over the enemy.

1.3 Commander's Tactical Display

The cost to remanufacture a Bradley A2 into an A3 will exceed $2.5 million per vehicle and will affect almost every aspect of the vehicle, from fire control to additional armor for protection. The training impact of all these changes will be significant. For the purpose of this paper, I will focus on how to train the Commander's Tactical Display (CTD), which provides digitization to the Bradley Commander in support of the Army's goal of Information Dominance. The CTD is an 8" by 10" flat panel display mounted directly in front of the Bradley Commander (BC). For input, a retractable keyboard is mounted below the screen and a thumb cursor for manipulating the cursor control is located on the commander's hand control station. The commander's hand control station provides trigger controls for the turret

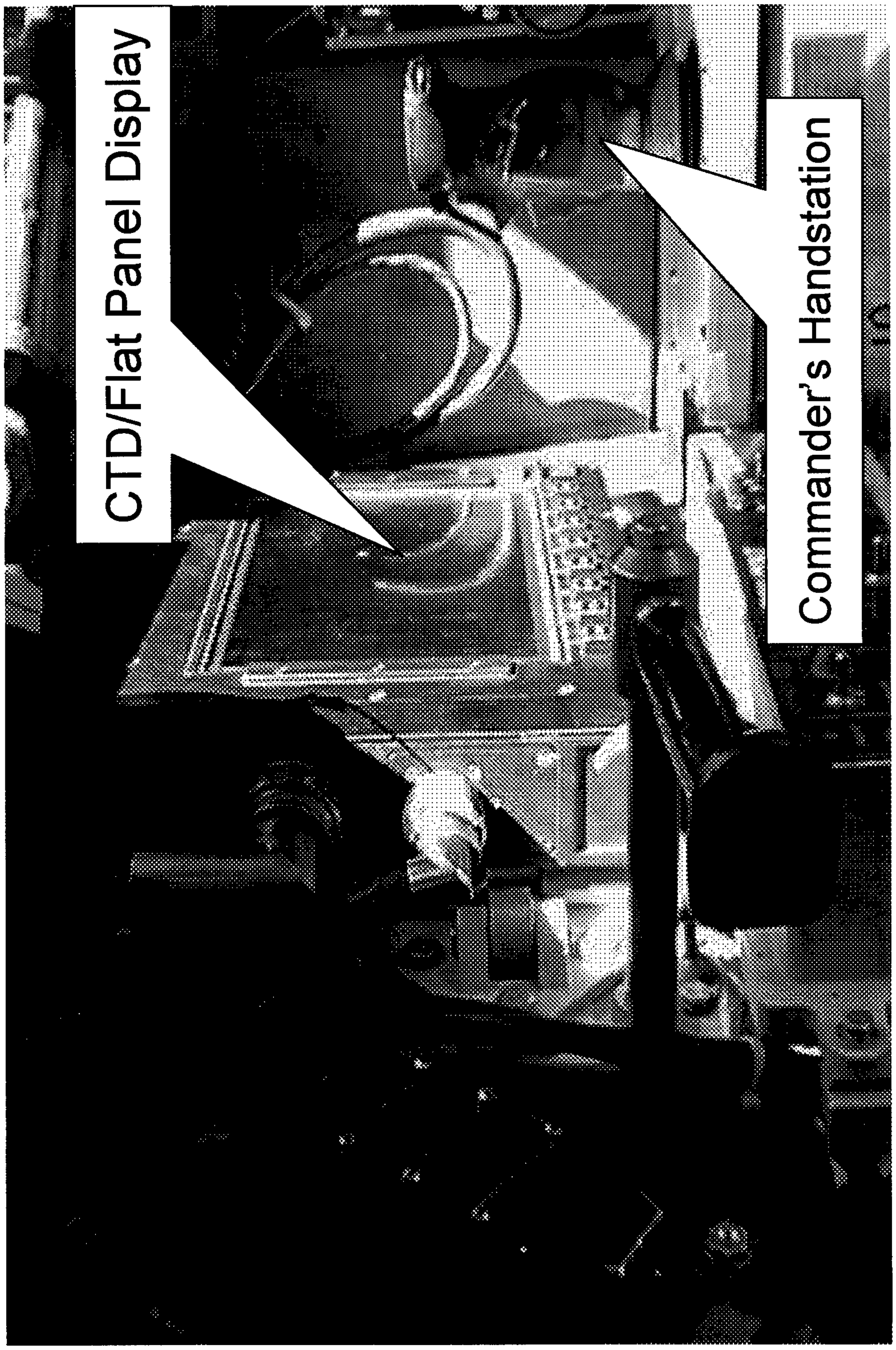


Figure 1. Commander’s Tactical Display

and weapon systems. Eight hard keys (with changing functionality) are also mounted on the bottom of the flat panel display. (See Figure 1. Commander's Tactical Display)

In order to correct the noted Desert Storm Deficiencies and to support the doctrinal priorities of the army, the CTD provides critical functionality on the soon to be fielded Bradley A3. While many of the changes to the Bradley A3 are merely improvements to existing features, the CTD is totally new and unfamiliar to soldiers in the field. Because of both its complexity and unfamiliarity, a significant level of training will be necessary to ensure proper use and to maximize its effectiveness at the individual level, which in turn will significantly impact both the organizational and operational levels. Unfortunately, there are no alternatives to training as the CTD is a unique device specific to the Bradley A3. While other military platforms will have similar displays to provide SA, exact compatibility is not possible due to differing roles on the battlefield. A recent Army initiative seeks to create a "common look and feel" with regard to Situational Awareness across all platforms. Regardless of the level of commonality achieved, the training issue may be simplified, but is not eliminated.

The Bradley Commander is the primary user of the CTD. Because of the display location in the vehicle's cramped turret, it is difficult for more than one person to view and interact with it, making it difficult to train digitization on the vehicle. Additionally, power/fuel requirements make training on the vehicle costly. (See Figure 2. CTD Screen)

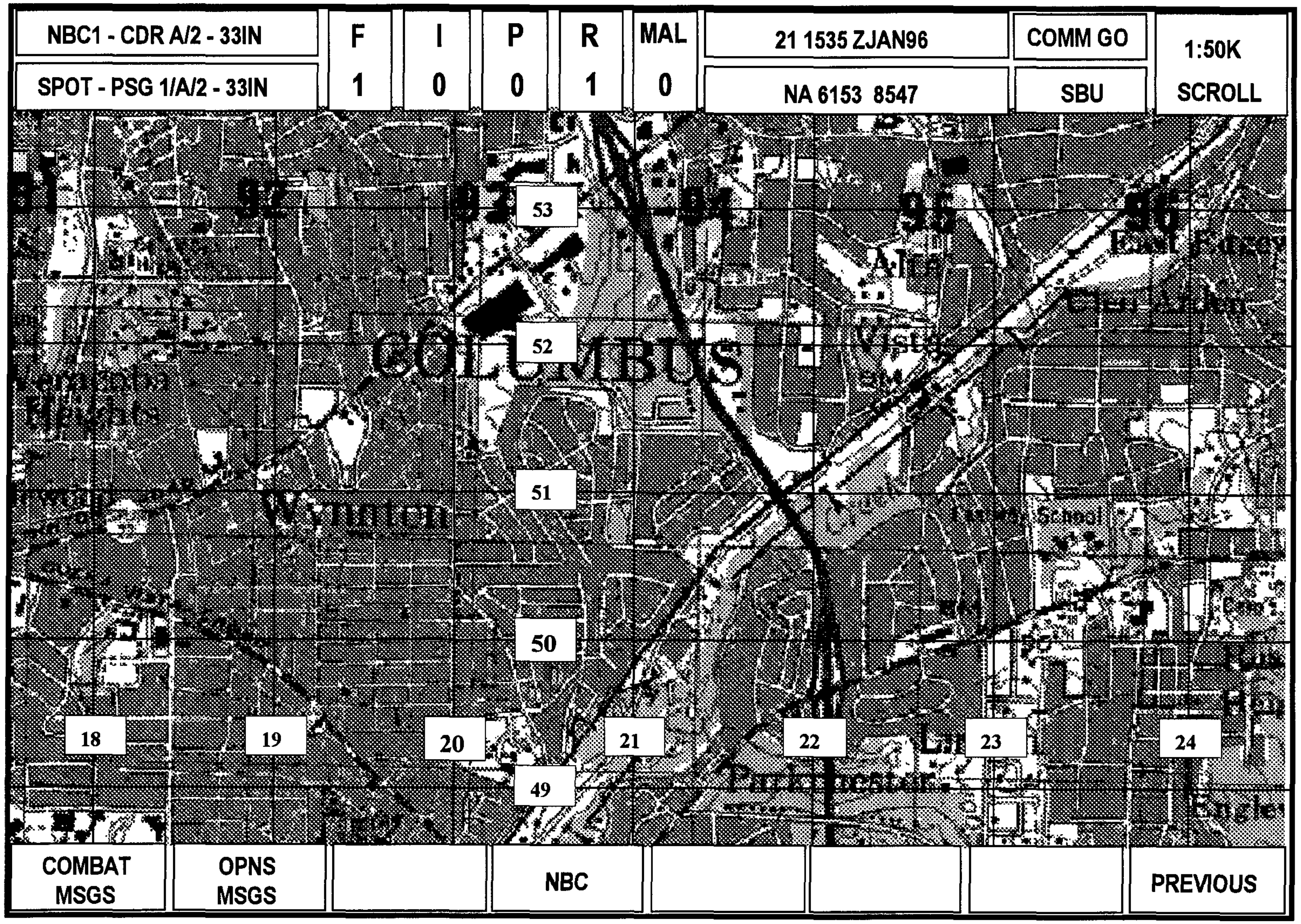


Figure 2. CTD Screen

1.4 Training Requirement

Due to the high cost of training on the vehicle, the US Army Infantry Center has stated the requirement for the Bradley Desktop Trainer (BDT). These requirements are captured in the Training Support Requirements to Bradley Fighting Vehicle System (BFVS): "a computer driven system that replicates the voice and digital communications capability of the actual Bradley system is required to help train and sustain the technical and perishable skills peculiar to the system..." (p 10). Additional guidance is found in the BFVS System Training Plan, which calls for "a digital communications trainer ... required to ensure that the crewmembers will achieve proficiency with the digital communications system" (p. 17).

The BDT is intended to provide a cost-effective method for training the Bradley Commander (BC) and his crew to efficiently use and interact with the CTD. Effective training at the individual level should improve both SA and C2 by providing a more accurate picture of the battlefield at all levels.

The screens as displayed to the BC have been developed to create an intuitive and user-friendly environment focused on the tasks most frequently used by a BC. The most difficult aspect will be learning to navigate between screens and different activities. Some tasks will occur every time the system is used. The most common and routine tasks have been deliberately placed at the beginning of the menu tree for ease and speed of use. Some tasks such as set up an address list distribution are conducted before enemy contact or during a lull in battle and consequently are not time sensitive. Other tasks such as creating combat reports are initiated during enemy contact concurrently with many other activities. Accordingly, these reports require speed and accuracy both to forward the information to appropriate decision-makers and to minimize the distraction of the soldiers creating the report. As a result, the Army must ensure its Bradley Commander's are trained for both timely and accurate use of the CTD.

United Defense, Limited Partners (UDLP), the Prime Contractor for the Bradley A3 Program is developing the A3 BDT. According to the BDT Student Guide, the primary purpose of the BDT system "is to provide a procedural trainer for learning and sustaining skills associated with the operation of the A3 CTD. Although other commander's station [devices] are visible, and partially functional, it (the BDT) is not intended to support a fully functional simulation of weapons systems." A BDT includes four workstations for trainees and an Instructor/Operator (I/O) Station for the primary trainer. Each of the four workstations simulates a CTD for an individual trainee. Figure 3 depicts the BDT Screen.

The instructor is responsible to teach the lesson plan, verbally generate scenarios, and monitor student progress. The simulator merely emulates the A3 CTD response to student interaction. The I/O station provides a visual representation of each of the four workstations to the instructor, who is responsible for many of those tasks automatically performed on an adaptive training system, such as providing instruction, generating scenarios, collecting data, and diagnosing student performance. In actuality, due to the complexity of these

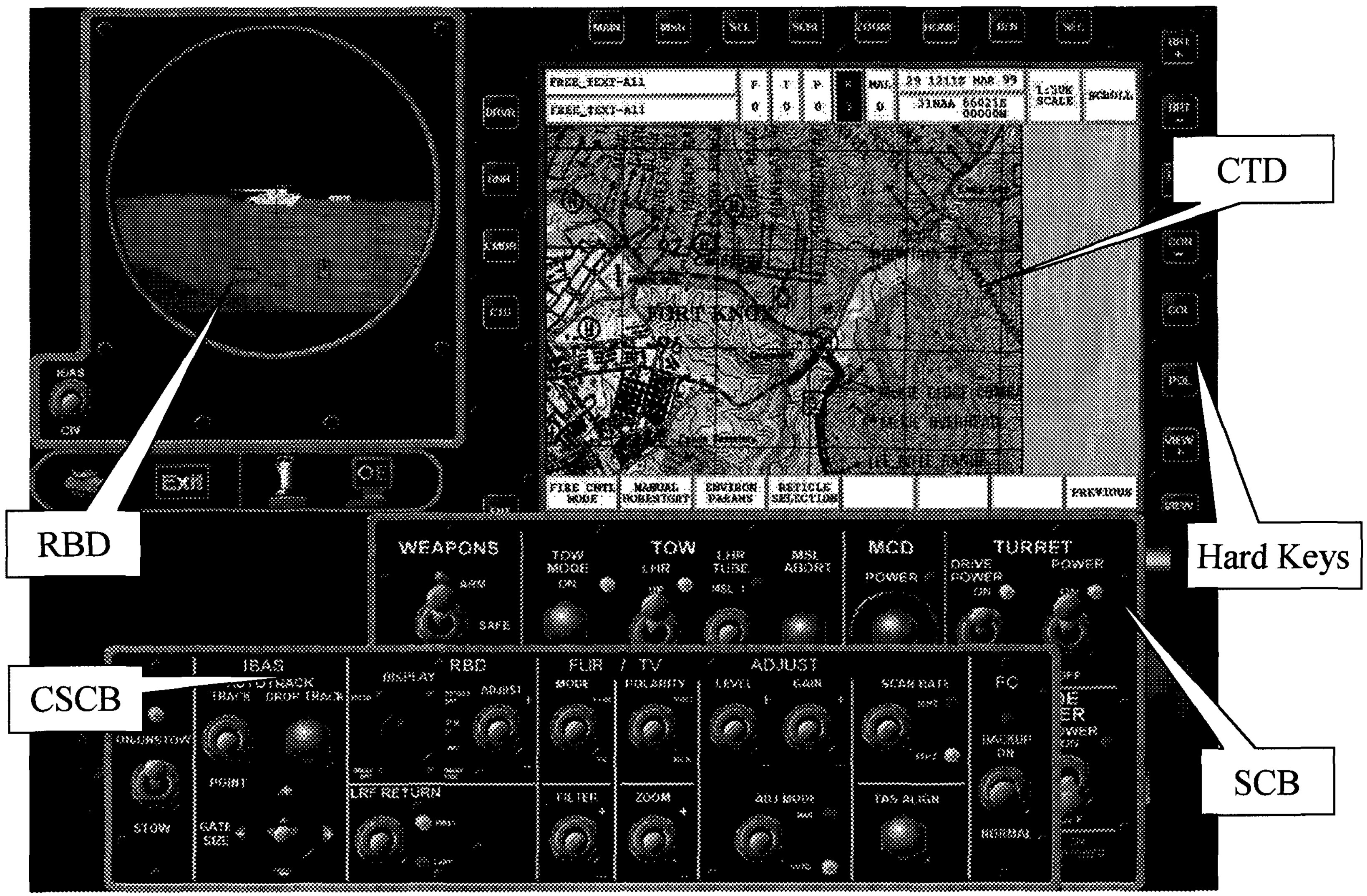


Figure 3. BDT Display

responsibilities, the instructor is able to little more than provide basic instruction, issue commands, and collect only the most rudimentary data with regard to a student's ability to accomplish a task. Similarly, all four students must work at the same pace without regard to ability. If one student is working at a slower pace, the others must slow down to his level to ensure the instructor can properly provide instruction, monitor performance, and collect training data. While this method might be acceptable for initial training on very basic skills, it severely limits the rate of training by requiring all students to work at the pace of the slowest student. This approach also makes sustainment training difficult because lessons cannot be tailored to the needs of individual soldiers who generally have different training requirements.

While the instructor is generally able to record pass/fail data based on student performance, he lacks the ability to accurately monitor the time required to accomplish these tasks. This leads to another limitation in the BDT system, because the time to perform a task is directly related to the individual's skill level. Soldiers should not be diagnosed as proficient if they cannot accomplish the task in a timely manner.

As a human computer interface device, the CTD is well suited to take advantage of embedded performance monitoring and an adaptive training strategy. Many of the aforementioned deficiencies could be easily resolved by adopting an adaptive training approach.

2 THE PROBLEM AND ITS SETTING

2.1 Problem Statement

> Traditional Knowledge Engineering methods have lead to a costly and inefficient development process for building the cognitive models necessary for Adaptive Training (AT) systems, making AT systems unaffordable to most organizations.
>
> This research aims to select and evaluate a cognitive analysis tool capable of developing the cognitive models necessary to provide an AT strategy for human-system interaction and embedded performance monitoring for procedural knowledge based systems. This experiment will focus on developing a set of cognitive models for training the Commander's Tactical Display (Human-System Interaction) on the Bradley A3. The intent is to determine the tool's ability to create consistent and accurate cognitive models by guiding domain experts through the knowledge acquisition process without the assistance of a knowledge engineer. To my knowledge, no other tools of a similar nature have been evaluated.

Well-developed cognitive models (knowledge bases) are necessary for the development of adaptive training and embedded monitoring systems. At present, the process to develop cognitive models is both costly and inefficient due to the high level of involvement required by a knowledge engineer who must essentially becomes a domain expert in order to adequately map and decompose tasks to an appropriate level. This high cost and inefficiency has led to the development of several automated cognitive analysis tools that attempt to guide the domain expert in the design of knowledge bases with minimal input from the knowledge engineer. The first step of this research involves reviewing the various methodologies and tools for designing cognitive models in order to select an appropriate tool for human-computer interaction tasks such as those associated with the Bradley A3 CTD. Once selected, the tool will be evaluated by comparing both the accuracy and consistency of selected models as created by various domain experts. If successful, we can be confident in the tools ability to successfully model the remaining BDT tasks necessary for the development of an effective Intelligent Training System that includes both embedded performance monitoring and embedded training.

The advantages of an adaptive training system for the military are immense. Such systems are more effective and efficient as they allow individuals to work at their own pace rather than to the slowest trainee. Additionally, an adaptive training system could very easily be hosted on the actual vehicle to make the training embedded. An embedded training system offers the following additional advantages; 1) performance monitoring of the soldiers in the field, 2) the ability to provide assistance as necessary, and 3) sustainment training to forward-deployed units that are away from home station for extended periods of time. This will become more critical as our military continues to participate in peacekeeping operations that frequently lead to decay in warfighting skills, as units are unable to train during these missions. Benefits of this adaptive approach potentially include:

- Increased training effectiveness based on individual assessment and requirements
- Better collection of diagnostic data to include accuracy and speed

- Increased transfer to an embedded training/monitoring system that can customize subsequent training sessions.

- Increased fidelity/training transfer

- Reduced long term cost due to increased rate of learning and removal Instructor/Operator equipment and personnel

An adaptive training strategy is expected to be more effective for either initial training or sustainment training. Because of the individualized self paced training, the rate of learning should increase. This would facilitate either the training of more soldiers on the same number of systems or an increased degree of training.

2.2 Research Question

> Can subject matter experts, without assistance from a knowledge engineer, consistently and accurately develop cognitive task models for their training domain? How can we leverage current knowledge acquisition tools and methods to develop a set of cognitive models that can enhance digital training on the Bradley A3 for both embedded performance monitoring and for an adaptive training curriculum development?

Is there a cognitive modeling tool capable of guiding domain experts in the development of knowledge bases with minimal assistance from a knowledge engineer? A major challenge in the development of intelligent tutors involves the tedious and costly input required by the knowledge engineer. Eliminating or even reducing the knowledge engineer's involvement can significantly reduce the overall cost in developing cognitive models. While the tool may not provide 100% commonality or accuracy between domain experts, resulting from semantic expression errors or an occasional missed step/method, an effective tool must provide a level of commonality that can significantly reduce the knowledge engineer's involvement.

2.3 Operationalized Research Question

> How can we extract expert knowledge from a group of subject matter experts to create a limited set of valid cognitive models for digital training on the Bradley A3?

A group of Bradley Subject Matter Experts (SME's) must be familiarized with knowledge acquisition methods and trained on the cognitive analysis tool. Subsequently, they must independently design a set of cognitive models for a set of tasks representative of the skills to be performed.

2.4 Research Purpose

> The purpose of this research is to validate an automated knowledge acquisition tool for procedural based tasks related to human-system operations as found on the Bradley A3.

Knowledge based systems have many potential advantages, from long term cost savings to more efficient and effective training. However, their development poses many challenges, which include organizing, representing, refining, and verifying the elicited domain knowledge. In general, this process has been time consuming and inefficient. Recently, the development of computer assisted knowledge acquisition tools have helped automate this process by eliciting information directly from the expert. Consequently, domain experts, with limited training in knowledge acquisition should be able to design cognitive models for a representative sample of selected tasks.

3 BODY OF KNOWLEDGE

3.1 Intelligent Tutoring

The current approach to training on the BDT is slow and inefficient at best. Soldiers must progress at the rate of the slowest student. During refresher training, all trainees must work on the same material rather than focusing on an individual area of weakness. A strong requirement for digital training exists that will enhance both initial and sustainment training and increase skill retention. William Sanders from the Army Research Institute found significant skill decay for digital procedural skills. More specifically, his research identified a 23%-52% (depending on the task) reduction in the number of soldiers able to meet the required performance criteria after 30 days without sustainment training. As COL Lynch, former commander of the EXFOR (the army's experimental brigade) has said, "I do know that these kids are losing the skill when they walk away from that ... box (a tactical display). I mean, they ramp up certain parts ... to be the best operator in the Brigade Combat Team. Well, you take him away from that box for about a week and now he's starting all over again. Not starting over, but he's lost a lot of those basic skills." (Interview w/ COL Lynch, 98) While the Bradley A3 is yet to be tested, we can assume a comparable decay of skills based on the similar nature of tasks being performed.

As an essential component of the Army's thrust toward Information Dominance, the CTD requires a training device that can improve soldier efficiency in order to maximize effectiveness of this valuable asset on the battlefield. With the need for a training device capable of both initial and refresher training the move toward an Intelligent Tutoring System (ITS) would be prudent. Experiments conducted by Williams, Reynolds, Carolan, Anglin and Shrestha, 1989; and Carolan, Williams, and Moskal, 1991; demonstrated dramatic improvement in performance when groups of individuals were trained using curriculum structured in accordance with the ideal student modeling principles embodied in intelligent tutoring systems compared to those trained employing curriculum generated by a conventional content analysis technique.

The BDT as currently designed represents a Computer-Assisted Instruction (CAI) system. A CAI system is limited by two applicable themes: 1) the learner is the focus of CAI, and 2) the computer is only the vehicle for instruction, not a method of instruction. (Steinberg, 1984) Therefore, neither the computer hardware nor software is able to determine the curriculum or lesson plan in a CAI system. CAI systems are also linear in nature, while an ITS is more holistic. If a student has difficulties with a particular task in a CAI system, the system merely loops back to a preceding step and resumes the instruction on the same path. By tracing the individual's performance, an ITS system is able to identify the subtask causing the difficulty and therefore adapt the curriculum to retrain that particular subtask. The ITS identifies and trains common tasks that may be required for several higher level tasks, thereby recognizing the relationship both across tasks as well as vertically within the task.

The aim of Intelligent Tutoring Systems is to develop tools and models that give the computer the capability of becoming more involved in the tutoring process. While the focus of an ITS is still the student, it turns the computer into a dynamic instructor rather than a static vehicle of information. As Brusilovsky (1995) explains, this is possible based on three types of knowledge: "knowledge of the subject matter, knowledge of the teaching strategy and methods, and knowledge of the student." The ITS (Figure 4) must monitor and diagnose the student by comparing his performance against the ideal student model. The ideal student model provides the detailed subject matter that allows the ITS to be adaptive in nature. Domain tasks are represented in the ideal student model as expressed by domain experts. The system can then compare the actual student performance against the ideal student model in order to diagnose and monitor performance. Based on this feedback the ITS uses the planner to develop an appropriate curriculum, specific to the student.

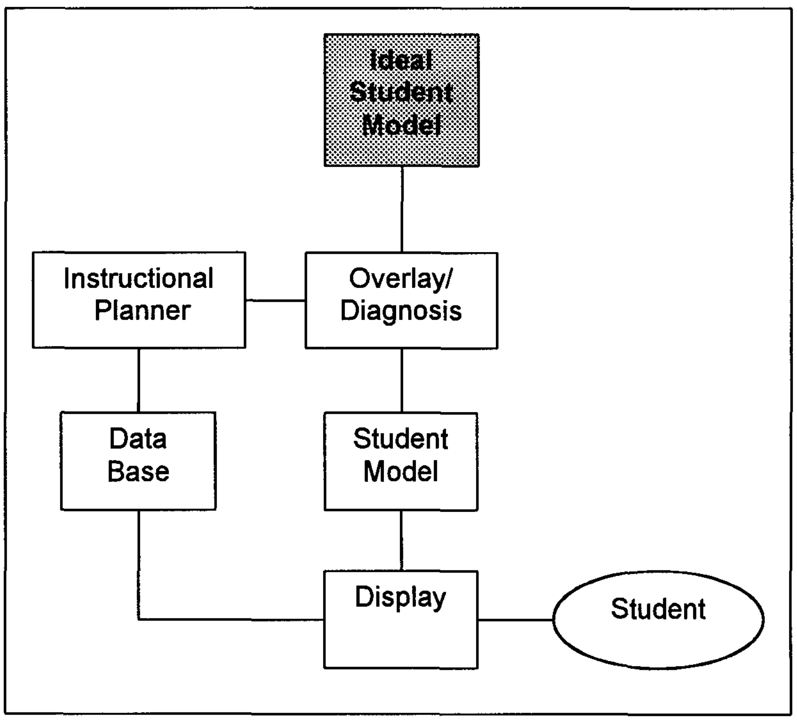


Figure 4. Intelligent Training System Model

Due to the dynamic nature of tutoring, the planner must be able to frequently modify the curriculum according to student performance. When a student makes a mistake, the ITS provides appropriate feedback so the student can identify his error and draw inferences from it. This use of the student model and ideal student models leads to a major advantage for ITSs. By "modeling user's learning methods and knowledge states, systems can adapt to their particular needs and abilities." (Dillenbourg, 1992) Because the ITS knows the student's knowledge state, it can provide more appropriate feedback suitable to the student's level of development and make more informed decisions regarding lesson content. Inappropriate feedback or assistance above or below a student's current level of development leads to frustration with the system during the learning process. A well-developed ideal student model eliminates this problem by directing feedback and subsequent instruction that is appropriate to the student's level of expertise. As a result, the instruction provided by the ITS is more relevant to the student's needs and weaknesses. The major disadvantage of CAI systems is their inability to adapt to student needs and weaknesses. By adapting the training to the individual student, a good ITS system can be expected to significantly increase the rate of learning.

3.2 Cognitive Models

As the basis for an ideal student model, cognitive models are generally composed of both declarative and procedural knowledge units. These knowledge units are associated with two types of memory; declarative memory and procedural memory. Declarative knowledge units represent facts about a specific domain or relationship between objects. For example, declarative knowledge units might include particular facts associated with a stoplight, like the significance of each color; red, green and yellow. A procedural knowledge unit is a group of facts (declarative knowledge units) in a specific pattern of association that is in turn linked to an action. In the example of the stoplight, an individual must correctly interpret the red light and then respond accordingly. Actions might include, judging the distance to the light, removing foot from the accelerator pedal and applying the appropriate pressure to the brake pedal to stop at the correct distance from the intersection or the car in front of him. As in the example, these actions can be observable physical behavioral actions or covert mental activities like making a decision, recalling something from memory, or adding or deleting information from working memory. As a skill is mastered through repetition, a task will become proceduralized, which means that it becomes automatic. There is very little if any interpretation; the individual just does the action automatically. These automatic procedural knowledge units are also referred to as production units or production rules. (Williams, 98) This procedural knowledge is typically the focus of computer interaction training like the CTD. The intent is to design a training system that will proceduralize the tasks and activities associated with the CTD so as to minimize the time the soldier spends interpreting information as he navigates the menu structure. This requirement for a procedural training system will play an important role in selecting an appropriate knowledge acquisition method.

3.3 Knowledge Acquisition Methods

Development of an effective ideal student model is imperative to the development of a functional ITS. While numerous knowledge acquisition methods exist, it is important to select an appropriate one for the development of the BDT ITS. Creating an ideal student model is essentially a knowledge engineering process that decomposes knowledge into basic units or production units that can be executed by a production system to simulate a cognitive task.

3.3.1 Methods

Numerous methods have been developed to facilitate the process of generating ideal student models. Generally, they fall into one of three categories: 1) machine learning, 2) machine aided, and 3) manual. Kotnour's (1992) research provides and excellent review and assessment of these various methods. In general, the knowledge engineer (KE) is heavily involved throughout the knowledge acquisition process making it both costly and inefficient. Ordinarily, the process might look like this; 1) KE Elicits knowledge from the expert, 2) KE organizes knowledge, 3) KE represents knowledge, 4) KE & expert refine knowledge, and 5) KE and expert verify knowledge. The focus of this research is to find a method that reduces the cost and inefficiency of the knowledge acquisition process by reducing this high level of involvement currently required by the KE. The following is a very brief review of Kotnour's findings, which conclude with the methodology that best supports the development of an adaptive trainer for the CTD and minimizes the active participation of the KE.

3.3.1.1 Machine Learning

While machine learning becomes more promising as technology advances, it has some significant drawbacks given current accumulation of knowledge bases and technology developments.

> The Machine-learning methods require very little direct interaction on the part of the expert or knowledge engineer. S/he is responsible for providing or gathering the data to be used by the machine-learning method. Machine-learning methods generate the knowledge from data. These methods automatically organize, represent, and refine the knowledge base. Machine learning is the epitome of knowledge acquisition. (Williams et al, 1993)

Kotnour concluded that these methods require significant domain knowledge prior to initiating the knowledge acquisition process. Since the purpose of the research is to develop a tool that minimizes the process of knowledge acquisition and direct involvement by the knowledge engineer, these machine-aided techniques were not considered. Unfortunately, these techniques require a significant base of domain knowledge or a large data set in order to generate a single knowledge unit.

3.3.1.2 Machine-Aided Methods

The second method of knowledge acquisition is Machine Aided as defined by Williams et al (1993):

> Machine-aided methods elicit, organize, represent, and refine the knowledge interactively with the expert. As the elicitation process proceeds, the machine implicitly organizes and represents the knowledge. Most machine-aided tools provide facilities to interactively refine the knowledge base. ... Machine-aided methods are typically automated versions of the manual methods. Automation aides in the organization, representation, and refinement phases.

Machine-aided methods are advantageous in that experts are able to directly transfer their knowledge independent of a knowledge engineer. These systems include automated knowledge-acquisition and knowledge-elicitation tools. In general, machine-aided techniques implement similar techniques to the manual methods, but use a computer to elicit much of the information. While some of the various methods are better than others, Kotnour (1992) eliminated most for several reasons. Some required too much prior domain knowledge, making them very specialized. Others focused excessively on the relationships between components and objects rather than on the procedural type knowledge associated with soldier tasks.

3.3.1.3 Manual Methods

Manual Methods are the most basic approach for designing ideal student models and are described as follows:

> Manual methods require the knowledge engineer (K.E.) to be directly involved in the complete process. The knowledge must first be elicited and then manually organized. Once the knowledge is represented, the KE must manually encode the knowledge in form acceptable to a specific knowledge-base system. Finally, the refinement and verification is done by a manual step-by-step examination of the knowledge base and the inferences derived from the knowledge base. (Williams et al, 1993)

With little aid from machines, the KE manually constructs the representation for input into a specific format of the software system's shell. In general, the KE uses the Interview, Observation, Interface Design, or Document Examination method to gather and structure his information. Of these techniques, the Interview method holds the most promise in extracting the data from the domain expert with minimal KE involvement. The Observation method relies heavily upon the KE's observations of the domain expert. Additionally, the KE must have a good grasp of the subject matter in order to sufficiently interpret the expert's actions.

The Interface Design method might be useful in the development stages of a system, but in the case of the Bradley A3, an interface design has already been implemented. Finally, the Document Examination process was similarly eliminated because of the heavy reliance upon the KE to wade through massive amounts of documentation in order to understand the knowledge domain and properly represent it. A brief overview of the interview method follows.

3.3.1.4 The Interview Method

Several interview methods are available for the KE to elicit information from the subject matter or domain expert. Examples include 20 questions and card sorting. In the first instance, the expert generates a series of 20 yes-no questions to diagnose a problem in the domain. Card sorting has the expert sort a series of cards provided by the KE. The cards are sorted by placing related concepts in a category based upon the expert's experience. Then the expert explains the interrelationships among the categories to the KE. Other examples include questionnaires, ordering trees from recall, hierarchical clustering and cognitive task analysis. (Williams et al, 1993.)

3.3.1.5 GOMS

In 1983, Card, Moran, and Newell developed the GOMS methodology for constructing psychological models of human computer interaction. (Kieras, 1985) A GOMS model has four components, Goals, Operators, Methods, and Selection Rules. The **G**oal is the state to be achieved, or end state. Goals may include subgoals. **O**perators are prerequisite tasks to achieve a goal. **M**ethods are a set or series of operators used to accomplish a specific goal. **S**election Rules are sets of discriminating conditions used to choose between different methods of achieving a particular goal. The GOMS methodology is essentially a process of decomposition in which the operators/steps needed to accomplish a goal are specified. These steps form a method. If more than one method of achieving the goal exists, then a selection rule is specified to determine the appropriate method based on the conditions. The process is further decomposed as the operators of a given method become subgoals, which are then broken down into operators, methods and selections rules. When all operators have been broken down to their lowest level (e.g. they cannot be broken down any more) the process stops. (Kotnour, 1992). Operators that have been fully decomposed are called primitive operators.

According to Card et al, (1983), all human computer interaction can be organized according to the GOMS methodology. Additionally, he proposed that a small finite set of primitive operators could define any human-computer interaction task. Consequently, the knowledge engineer can predict the time to execute a task without having to collect execution time, by using the GOMS methodology and the relatively small set of well-defined primitive operators. After evaluating approximately 40 knowledge acquisition methods, Williams et al (1993) concluded that the GOMS cognitive task analysis technique to be the most effective for procedural type tasks.

3.3.1.6 CAT

In 1992, Kotnour evaluated the Cognitive Analysis Tool (CAT); an automated knowledge-acquisition tool based on the GOMS methodology. Intended for use by the domain expert, not the knowledge engineer, CAT provides an interface that attempts to walk the domain expert through the GOMS process by employing a structured interview process. The expert specifies information as prompted by a series of dialog boxes. CAT employs the following dialog boxes to guide the expert through the knowledge-acquisition process: top-level goal; used to define the top-level goal; method editor dialog box, used to define the set of steps to accomplish a goal; selection rule editor dialog box, used to define the selection rules for alternative methods; and yes-no dialog boxes, used to ask the expert if alternative methods or selection rules exist and if a goal is primitive or not.

CAT begins by requesting the top-level goal. Next, the expert is prompted to specify a method that can be used to accomplish the goal. The method consists of a series of steps required to meet the goal. Upon completion of the method, the expert is requested to list any alternative methods to achieving the goal. After all the alternative methods and steps have been specified, the expert is asked to provide the selection rule or rules that specify when to use which method to achieve the goal. Having completed the top-level goal, methods(s), alternative(s) and selection rule(s), CAT converts each step into a subgoal, and the process begins over again beginning with the method for each subgoal. This process of converting steps into subgoals to be further decomposed continues until the steps for each method are considered primitive. A primitive step is identified when it can no longer be decomposed.

Kotnour's research, in 1992, tested CAT by having 40 subjects create cognitive models of the Apple's "cut and paste" feature. His research found that subject models generated by CAT were 72.8% accurate when compared with baseline, as defined by Kieras (1988). When considering the only physical or explicit actions, the accuracy increased to 82.7%. In general, the inaccuracy resulted from an absence of cognitive or mental primitives generated by the test subjects. The subjects specified only 28.3% of the mental primitives from the baseline. Mental primitives as described by Kieras et al are "non-observed and hypothetical, inferred by the theorist or analyst." As non-behavioral steps taken to accomplish the task, subjects or experts generally implicitly perform them. Consequently, since the test subjects were not cognitive analysts it is unlikely that they would include mental primitives for implicit actions such as "determine position of beginning of text." Other examples of even more cognitive-based primitives include "recall" and "store in long-term memory" or "store in short term memory" as prescribed by Kieras.

A major concern of Kotnour's (1992) in evaluating the accuracy of automated aids for generating knowledge bases is the question of how one defines accuracy. In Kotnour's experiment, he defined accuracy relative to a baseline model as created by Dr. David Kieras, an authority in cognitive analysis. In general, the subject models left out specific primitives as specified by Kieras. However, in some instances, legitimate steps were added. In general, the models were consistent in higher-level tasks but differed in the level of detail they achieved. Kotnour concluded that the "accuracy of the model depends upon the use to which it will be put. If the model is to be used to stimulate a cognitive process in terms of execution time, a highly-detailed definition of all mental primitives or operators is required." On the other hand, the reference to primitive mental tasks might be confusing if one were describing a task to a non-cognitive psychologist. For the purpose of this research in developing models for training simulators, the level of required accuracy is very high. An effective training simulator not only monitors performance by mapping tasks to the ideal student model, but it also identifies the level of proficiency as indicated by the speed with which he accomplishes the task. Consequently, the accuracy of primitive cognitive tasks become more important by allowing the knowledge engineer to accurately predict time required to learn and to execute the task.

Kotnour also evaluated consistency between subject models and found a very high level of consistency of 88.5%. That is, the models contained 88.5% of the same primitive steps for any given primitive method. Because of the high consistency, we can infer that an accurate knowledge base can be developed in a domain that does not have a baseline model for comparison, like the BDT. Inconsistencies were generally found at the primitive task level. While most subjects managed to properly model the task, they may have omitted certain primitive tasks by assuming they were understood by their higher-level subgoal. For example, many subjects might assume that the task push a button includes the task release the button, thereby omitting it from their task list. As the subjects were not knowledge engineers, they should not be expected to decompose the task into its most primitive level, especially when lower level cognitive tasks are involved such as discriminate between two options. Based on these findings, CAT-HCI was developed to correct these deficiencies by prompting the domain expert to capture lower level cognitive tasks.

3.3.1.7 CAT-HCI (Cognitive Analysis Tool – Human Computer Interaction)

CAT-HCI is based upon improvements identified by Kotnour's evaluation of the CAT (Cognitive Analysis Tool). It guides domain experts through the knowledge acquisition process resulting in a model that predicts the time to execute the interactions described as well as the degree of consistency between interactions described for a particular human system interaction. CAT-HCI was improved to overcome deficiencies noted in CAT, specifically, the level of detail required by the user such as primitive mental operations. The test subjects consistently omitted steps such as "Determine Position of Beginning or End of Text or Verify" operations. While the subjects generally implied these primitive mental operations, their omission has a severe impact for making predictions concerning the time to execute specific tasks from the cognitive model. Consequently, CAT-HCI is designed to direct the domain expert towards the primitive operators that may make-up a step. This is accomplished by prompting the user to select from a pull-down menu of possible subsystem operations, which include; Arm-Hand-Finger Operations, Visual Operations, Auditory Operations, Cognitive Operations or Motor Speech Operations. The user follows the menu tree to select the primitive operator or sequence of operators necessary to execute the step. Once selected the step or series of steps are automatically generated and inserted into the method description. Finally, the user is prompted to select the source of information responsible for triggering the step selection. The result is a precise cognitive model that can accurately predict both the detailed sequence of steps and the time required to accomplish a particular task. (Williams, 98) Figures 5-7 provide examples of how CAT HCI dialogue boxes and pull down menus are used to select a primitive operator.

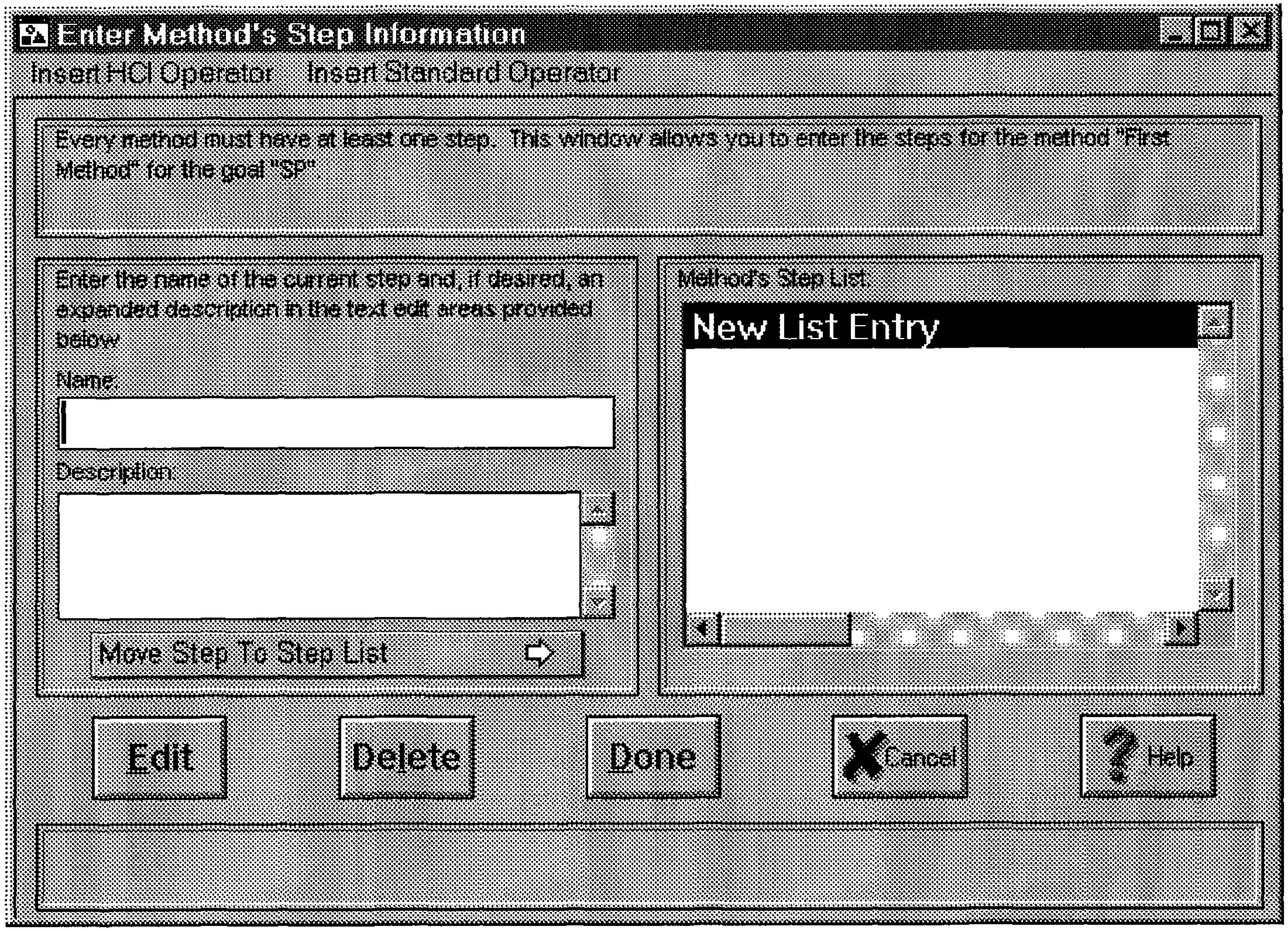


Figure 5. CAT-HCI Guidance Mode (Requesting a new step)

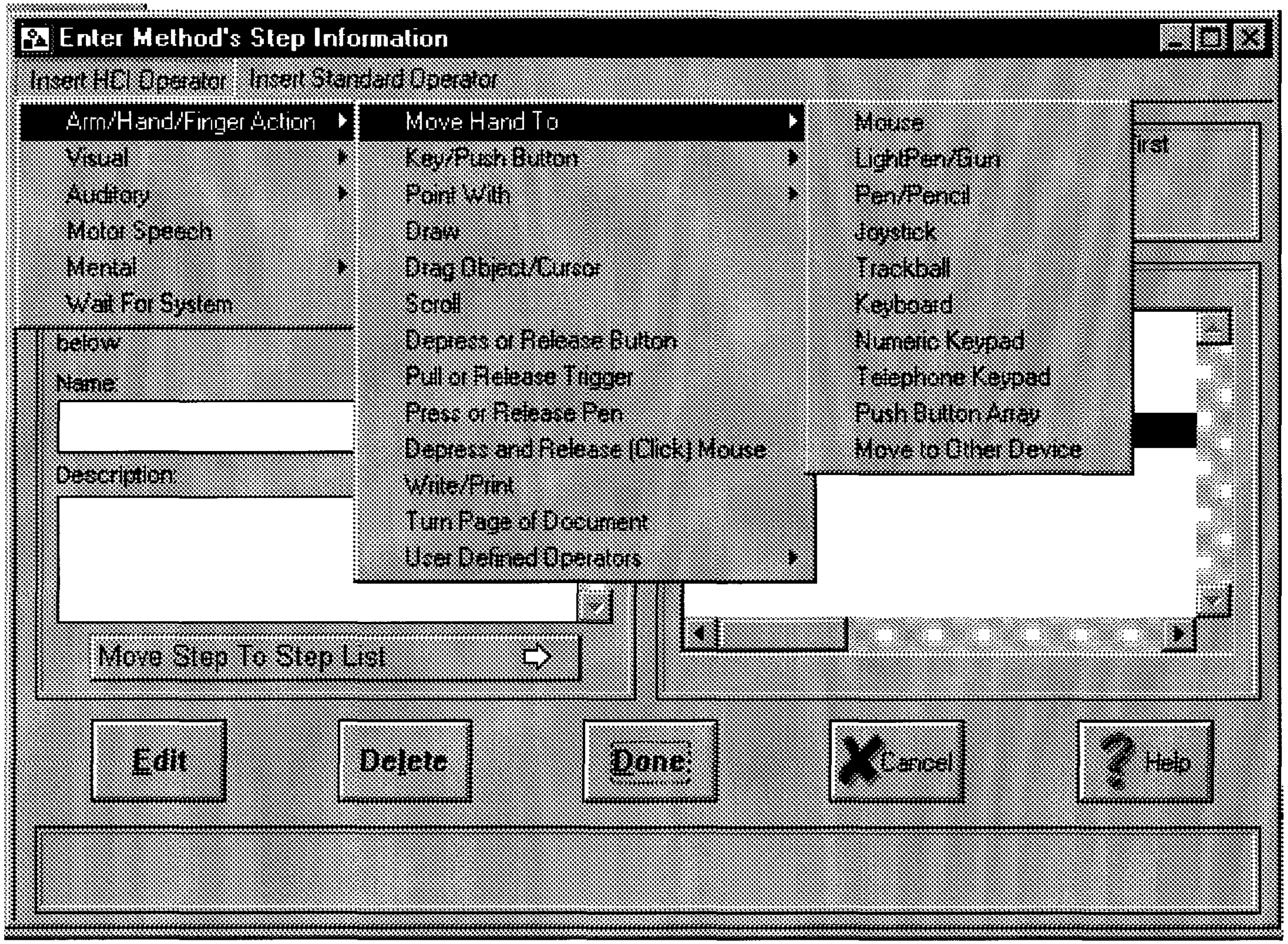


Figure 6. CAT-HCI Pull Down Menus (Used to Select *Move Hand to Joystick*)

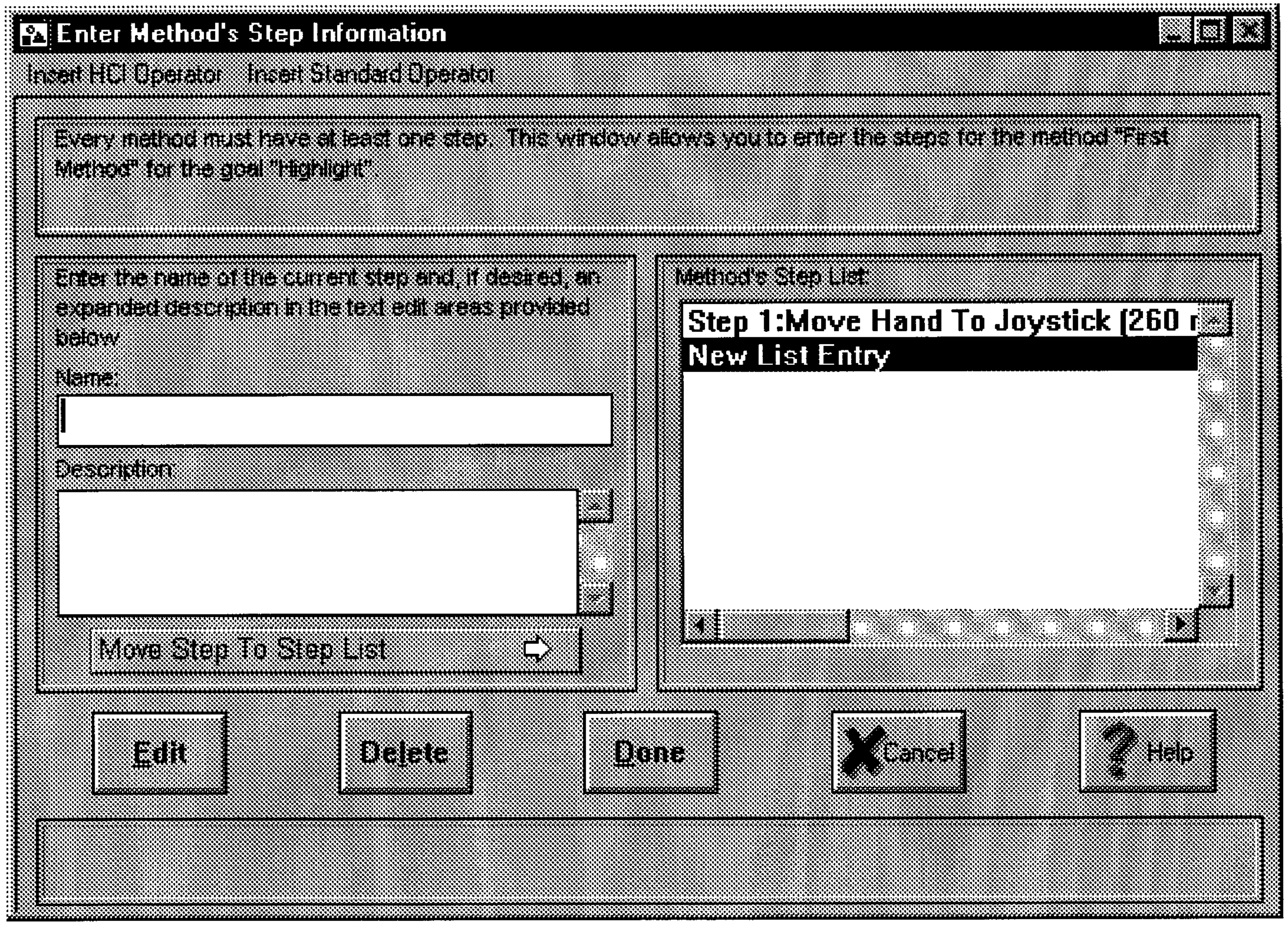


Figure 7. Primitive Operator Automatically Inserted by CAT-HCI (Notice 260 ms to complete task is included)

4 TESTING AND EVALUATION METHODOLOGY

This chapter details the process and procedures employed in testing and evaluating the developed knowledge acquisition tool called Cognitive Analysis-Tool Human Computer interaction (CAT-HCI). The operational purpose of this research is to develop a set of cognitive models for the Bradley A3 CTD. This set of tasks can be used to demonstrate potential future capability given appropriate funding and time. More importantly, the research purpose is to validate the CAT-HCI tool so that it can be relied upon in the development of subsequent models, not only for the Bradley A3, but also for other procedural human-system interaction tasks. By guiding the user through a process that describes human computer interactions, an engineering model is developed that predicts the time to execute interactions as well as the degree of consistency between interactions. This experiment evaluates CAT-HCI by having 18 domain experts use the tool to develop four cognitive models for a set of Bradley A3 digital methods. A minimal variance between these subject models and a high level of correspondence between the baseline model and subject models as created by different experts demonstrates the accuracy and consistency with which cognitive models of human-systems interaction can be generated by individuals not skilled in the cognitive task analysis approach. If the tool can generate accurate and consistent models, then by analogy, the inference can be made that the tool can be used to generate cognitive models in other human-system interaction tasks.

The purpose of the evaluation methodology is to investigate CAT-HCI's performance relative to meeting the research objective. In other words, can the automated knowledge acquisition tool guide domain experts through the GOMS process to formulate consistent and accurate cognitive models for eventual implementation by software engineers? Consistency between the different subject models and accuracy between subject models and the baseline should determine the effectiveness of the tool.

4.1 Question to Be Answered

How accurate and consistent are the CAT-HCI generated cognitive models as developed by domain experts with minimal knowledge acquisition training?

4.2 Experimental Plan

The methodology involves 1) determining accuracy by comparing subject model operators and predicted times against those of the baseline models, and 2) determining consistency by comparing subject models to each other. Accuracy is the level of agreement between a subject model and its respective baseline model. Consistency is a measure of the level of agreement between subject models. The tool must be able to generate accurate and consistent models from different experts in the same domain. (Figure 6. Conceptual Model)

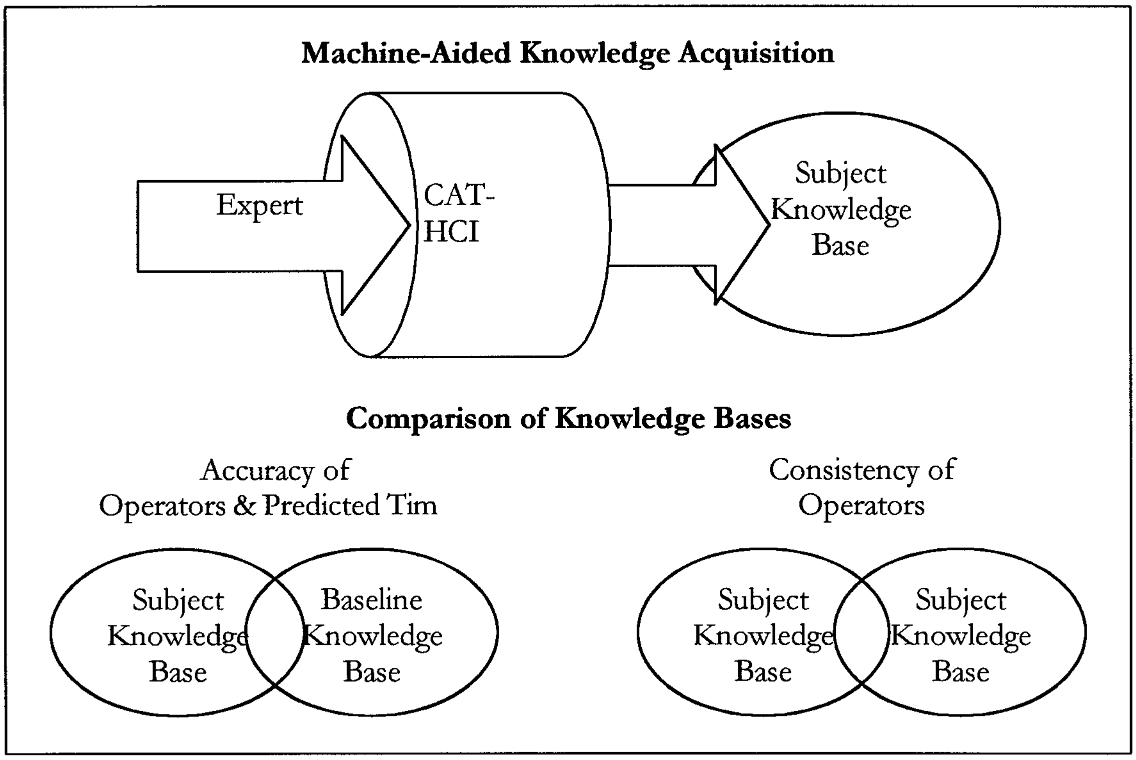


Figure 8. Conceptual Model of Methodology

4.2.1 Possible Sources of Variance for the Experiment

This section describes the possible sources of variance for the experiment. Each source is described in terms of how it could influence the experiment and how it will be controlled and accounted for in the experiment.

4.2.1.1 Type of Experimental Task.

The type of task is a source of variance because the task type might influence the ability to elicit knowledge. By choosing well-structured tasks, such as those defined in 4.5 Task Definition, this type of variance is controlled. While both well-structured and ill-structured tasks can be used for the elicitation of knowledge, well-structured tasks are much better suited to evaluate knowledge bases as generated by the tool and by extension, the tool itself. A well-structured task must be completed by a clearly defined set of actions, allowing accuracy of the models to be calculated. An ill-structured task may not have any one correct knowledge base, making it difficult, if not impossible, to measure the accuracy of the model bases.

4.2.1.2 Subjects used as experts.

The subjects used as experts for this experiment are an obvious source of variance. Because the Bradley A3 is still in prototype phase, the pool of Subject Matter Experts (SME) is very small and is composed almost entirely of the 18 subjects in this experiment. Additionally, as discussed in 4.3 Experimental Subjects, the SMEs are a relatively homogeneous group. They are all male, they have a high concentration of military service, and they compose the training base for the Bradley A3. Additionally, each subject was screened on his ability to perform the task.

4.2.1.3 Experimental Procedure

The experimental procedure is also a source of variance and therefore, was designed to minimize any variation in the conduct of the experiment. The procedures used in each session are detailed in Appendices C-E. Each step of the experiment was performed for each subject with any deviations noted.

4.3 Experimental Subjects

4.3.1 Subjects – Who Are They

As the system is yet to be fielded, the pool of domain experts remains very small and consists of approximately 20. These experts consist of 10 UDLP contractors responsible for training the trainers and newly equipped units as well as 5 members of the Bradley A3 NETT (New Equipment Training Team) which is headquartered at Fort Benning, Georgia. An additional 4-6 Subject matter experts work in the Bradley Proponent Office which has been critical in representing the user during design and production of the vehicle. The Bradley A3 underwent extensive train-up and testing at Fort Hood, Texas between 15 September and 15 November 1999. Consequently, many of the SMEs were concentrated at Fort Hood for this event. I coordinated with the Infantry Center to take advantage of this condition in order to conduct the experiment during the Bradley A3 Test and Evaluation. Initial data collection from the 10 UDLP contractor domain experts was conducted at Fort Hood, TX from 1-4 November 1999. I collected data from an additional four military SMEs at the UDLP facilities in Orlando on 2 December 1999 while they were supporting the ITSEC conference. I traveled to Fort Benning, GA to collect data from the final four subjects on 18 January 2000.

4.3.2 Sample Size

Because of the small pool of SMEs, the sample consisted of 18 subjects, which is nearly the entire pool.

4.3.3 Criteria for Selection

These subjects are being selected based on their expertise on the system. They either are all instructors for the Bradley A3 or were instrumental in the development of the system. This is an extremely homogeneous group, consisting of male active duty or retired soldiers. The 10 UDLP contractors consist of retired Non-commissioned officers with at least 20 years of Infantry or Cavalry Experience. The remaining subjects were active duty infantry soldiers with at least 6 years of service, with most having more than 15 years service. As an experimental control, each subject was required to perform the tasks on the vehicle or BDT prior to the experiment.

4.4 Experimental Apparatus

Each subject was screened on the task described in section 4.5, Tasks to be Modeled. Subjects conducted the screening task on the Bradley Desktop Trainer (BDT) or the Bradley A3 emulation software installed on an IBM compatible workstation. As previously mentioned (section 1.1.4), the BDT is a procedural trainer for learning and sustaining the skills associated with the Bradley A3 Commander's Tactical Display (CTD). The BDT emulates the CTD in order to train soldiers to create, send, receive and read various battlefield reports and to master the skills associated with situational awareness and digitization. UDLP kept four BDT workstations at Fort Hood to support this experiment. Four other BDT workstations were located at the UDLP Facility here in Orlando for the second group of subjects. As there were no BDT workstations available at Fort Benning, I installed the Bradley A3 emulation Software on four PC workstations to conduct the experiment. Each BDT workstation consisted of:

- Pentium II 400MHZ Computer
- High resolution 20" touch screen monitor
- Commander's Data Entry Device (CDET) – "Tactical Keyboard"
- Commercial Off the Shelf (COTS) Keyboard
- Commander's Hand Station (CHS)
- Mouse
- Speakers

The cognitive analysis tool CAT-HCI version 95 was loaded onto the BDT workstations so that up to four subjects could participate simultaneously. An additional benefit of using the BDT workstations to conduct the cognitive analysis is that the subjects had convenient access to the training device as a point of reference throughout the process.

4.5 Tasks to Be Modeled

There are numerous tasks associated with the CTD, from Maintenance and Logistics to Command and Control. Due to current software immaturity, only a small subset of the total tasks is currently available. See Appendix A for complete list of currently available of tasks. For the purpose of this research, I have selected *Create SPOT Report* because it is a routine report that will prove critical on the digital battlefield. It is not overly complex in nature, nor does it require the soldier to wait on the system or an external response to accomplish the task. The SPOT Report is commonly used on the battlefield to provide a brief update on enemy sighting. It requires mental, perceptual and physical primitives and includes a number of lower level methods suitable for modeling and analysis. The lower level methods that will be modeled include 1) Enemy Size, 2) Enemy Activity, 3) Location, and 4) Enemy Unit. Screen design templates and the baseline cognitive models as designed by the author are outlined in Appendix B. There are multiple techniques for entering and manipulating data (the thumb cursor control, hard keys and keyboard) on the CTD. During the experiment, the cursor control technique was directed to reduce the number of alternative methods.

4.6 Experimental Procedure

The experimental session consisted of three phases: 1) Subject orientation and screening test, 2) CAT-HCI familiarization, 3) CAT-HCI knowledge acquisition. Phase one was used to explain the experiment and verify each subject's expertise of the tasks to be modeled. Phase two was necessary to orient the subject to the CAT-HCI software so that he could adequately use it to model the tasks. The third phase involved the interaction of experimental subjects with the tool. Figure 7 contains the overall flowchart used for the experiment. Figure 8 shows the detailed flowchart for processing subjects. Multiple subjects participated in each of the five experimental sessions that lasted between six and eight hours. Groups of four subjects participated in four of the sessions while two subjects participated in a fifth session for a total of 18 subjects.

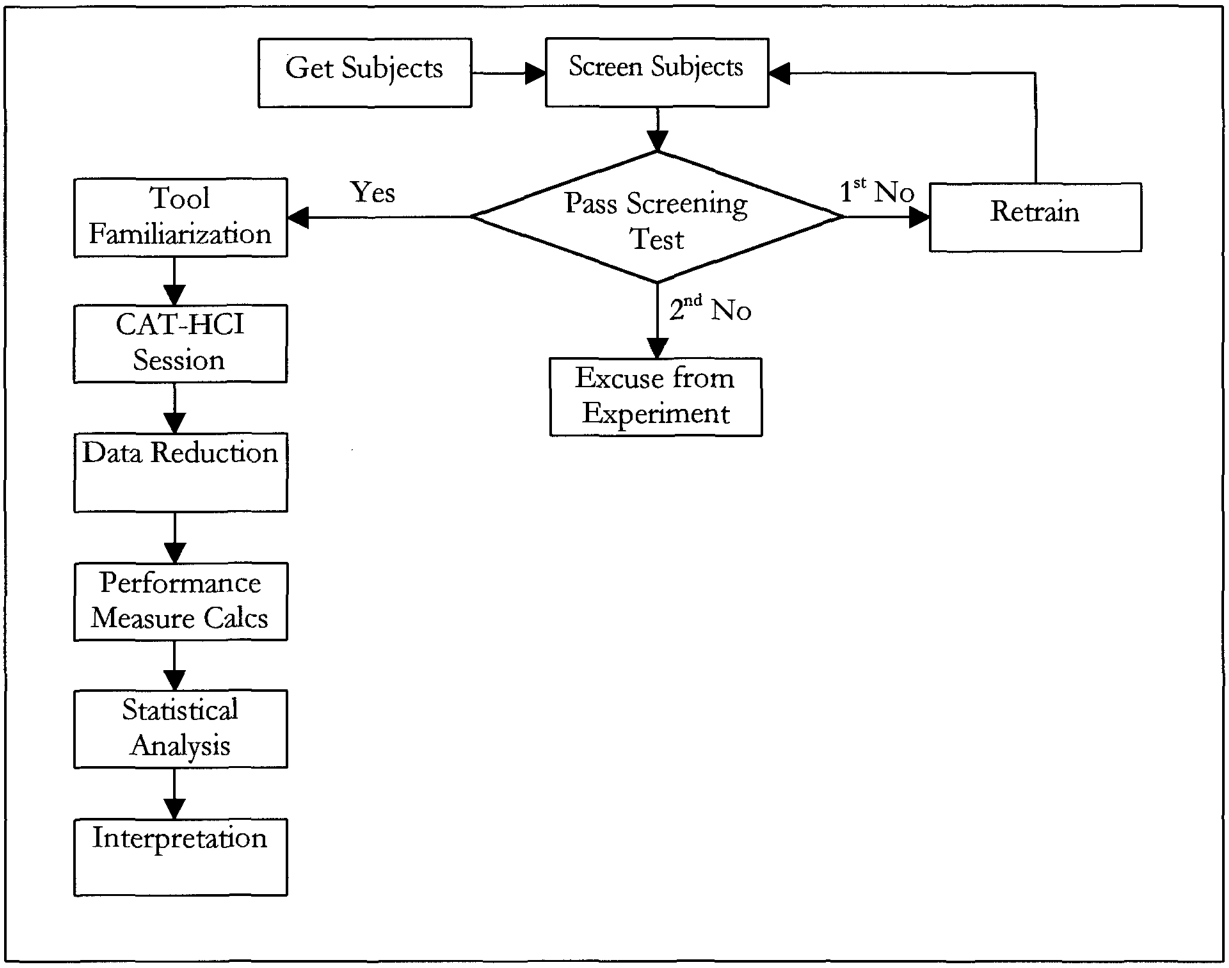


Figure 9. Flowchart for Experimental Procedure

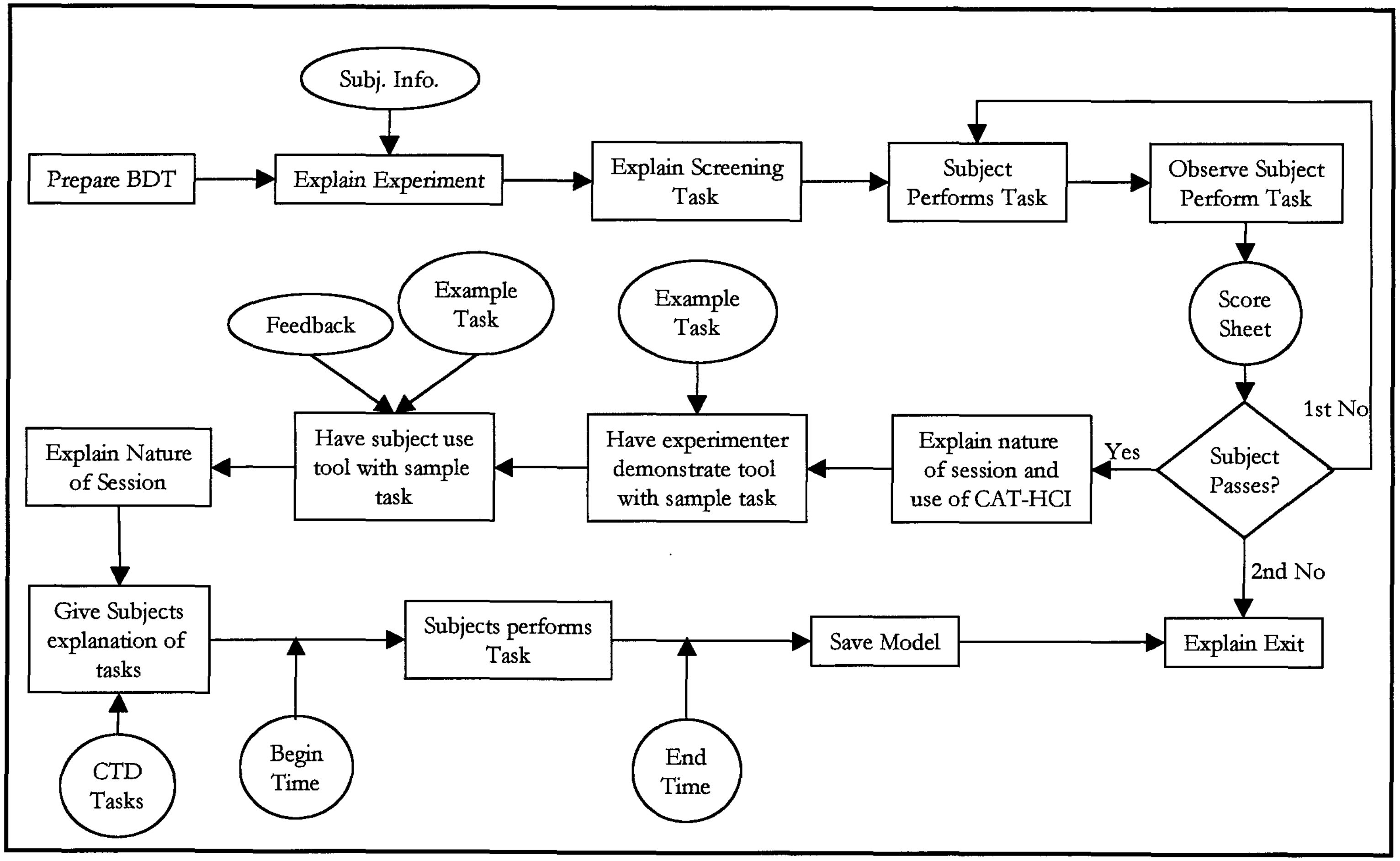


Figure 10. Detailed flowchart used for processing subjects

4.6.1 Subject Orientation and Screening Test

The purpose of this phase was to familiarize the subjects with the experimental process, to screen the test subjects relative to their level of expertise, and to collect personal data on each subject. While all the subjects included in the sample were considered to be experts in the A3 domain, their expertise in the operation of the CTD varied. This screening test ensured that each subject was proficient on the tasks to be modeled. If a subject could not perform the task correctly the first time, he would be allowed to review the task and try again. Subjects having difficulty a second time would be asked to leave the experiment. The performance measures were based on The Test Plan for the Bradley Desktop Trainer Version.1 (15 March 1999). The scoring matrix for this session can be found in Appendix C. This phase generally lasted about 30 minutes, and included about 15 minutes for the explanation of the experiment, 10 minutes for the collection of personal data, and 5 minutes for the actual screening test. All subjects successfully completed the screening phase the first time.

4.6.2 CAT-HCI Familiarization Session

During the CAT-HCI familiarization session, which lasted about 5 hours, I familiarized the test subjects with knowledge acquisition concepts and explained the functions and command tools of the CAT-HCI interface. The phase included three parts: 1) explanation of the tool by the experimenter, 2) guided examples completed by the subjects as explained by the experimenter, and 3) subject use of the tool to create a sample model with feedback from the experimenter. The purpose of the session was to train the subjects on the use of CAT-HCI so that they would be proficient with the application during the knowledge acquisition session. Appendix D contains an explanation of the application. Unlike its predecessor, CAT-HCI provides a finite set of primitive operators, which are deliberately elicited from the subjects. Because the BDT was specifically designed for the Bradley A3, it has some unique interface characteristics when compared to a more traditional computer workstation. For example, the cursor control device is mounted on the commander's hand station and is manipulated by the Bradley commander's thumb. CAT-HCI does not have a predefined primitive operator for manipulating a thumb cursor control. Because is most closely resembles a joystick in its operation, subjects were directed to use joystick primitives when referring to the thumb cursor control device. Other primitives as shown in Table 1 have been defined in advance to reduce ambiguity of terminology.

Table 1

Defined BDT Primitives

	Action	Use CAT HCI Primitive
Physical Primitive	Press Hard Keys	Press button
	Point, using thumb cursor	Point, using joy stick
	Press thumb cursor	Press button
Mental Primitives	Find Specific Icon/object	Look for object in viewing space

4.6.3 CAT-HCI Knowledge Acquisition Session

Subjects used the CAT-HCI tool to create the models for the methods defined in section 4.5 Tasks to Be Modeled. The task explanation for this session is given in Appendix E. This session lasted between 90 and 120 minutes based on individual subject proficiency.

4.6.4 Pilot Test

I conducted a pilot test on 29 October in the Industrial Engineering (IE) Computer Lab. Four military graduate students from the UCF Industrial Engineering Department volunteered to act as test subjects. Since the BDT workstations were unavailable, I used the A3 CTD software emulation package to train and screen the subjects. Both the A3 CTD emulation software and the CAT-HCI software were loaded onto four workstations in the IE computer lab. Unlike the actual test subjects, the pilot test subjects were not familiar with A3 CTD. Therefore, I began the pilot test by familiarizing the pilot test subjects with the CTD. Once the subjects were comfortable with the SPOT report and NBC reports, I ran all four subjects simultaneously through the experimental process from the screening test to the completion of the models. The pilot test provided a great opportunity get a better feel for the experimental flow. Probably the most significant lesson was the length of time necessary to train individuals on the use of CAT-HCI. The CAT-HCI familiarization process took more than two hours for the pilot test subjects and could be expected to take even longer for the experimental subjects due to their unfamiliarity with knowledge engineering. The pilot test subjects were all military officers in the UCF Training Simulation graduate program and were therefore much more familiar with knowledge engineering than the experimental subjects.

4.7 Data Collection and Reduction

Data collection began by extracting the primitive steps from the models developed by each subject. The primitives from each method were compared between subjects and the baseline to ensure semantic agreement. Each primitive method was composed of sequentially ordered primitive operators. All primitive operators came from a finite set as defined by CAT-HCI for human computer interaction and included an expected time for completion. In order to facilitate a more accurate comparison, in some cases the primitive operators from subject models were adjusted to the reflect agreement with the baseline version if the primitive had essentially the same meaning or intent. For example, to look for something on the screen, most domain experts correctly selected the primitive *Look for Object in Viewing Space,* while others selected *Look at Area in Viewing Space* or *Look for Location in Viewing Space.* As each of these primitives have essentially the same intent, the subject was given credit for agreeing with the baseline primitive of *Look for Object in Viewing Space.* After verifying subject models for correct terminology, the models were grouped according to each of the four major tasks, Enemy Size, Enemy Activity, Enemy Location, and Enemy Unit for comparison.

In order to compare between subject models and the baseline, another problem must be resolved as not all subjects arrive at exactly the same solution, nor do they necessarily have the same number of steps. By omitting or adding primitive steps in relation to the baseline, subjects frequently arrived at a sequence of operators that differ in length, making it difficult to compare across subjects. Card, Moran, and Newell (p.190) resolve this problem by placing the sequence of operators into correspondence and then assigning a value to show how well they match. For example, the baseline sequential acronyms for operators making up the method Highlight are: PLFOVS, CLFOVS, IHGT, CPWJ, MPWJ, CPB, MPB. The following algorithm accounts for both the predicted sequence and the observed sequence in order to provide a basis of comparison.

Table 2

Unadjusted Comparison of Predicted and Observed Steps

Predicted:	PLFOVS	CLFOVS	DS	CPWJ	MPWJ	CPB	MPB
Subject 3:	PLFOVS	CLFOVS	CPWJ	MPWJ	CPB	MPB	

Notice that for Subject 3, the sequence is correct, but the primitive DS has been omitted, resulting in a lack of correspondence between the remaining three operators. The solution is to insert a dummy (X) operator to realign them so they correspond.

Table 3

Adjusted Comparison of Predicted and Observed Steps

Predicted:	PLFOVS	CLFOVS	DS	CPWJ	MPWJ	CPB	MPB
Subject 3:	PLFOVS	CLFOVS	**X**	CPWJ	MPWJ	CPB	MPB

Notice before the algorithm was applied; there are only two matches out of a possible seven, resulting in 29% accuracy. By inserting the dummy operator, the number of matches increases to six out of seven for an accuracy of 85%, which much more accurately reflects the similarity between the two.

Another issue relates to the method *verify*, which is a required method for each of the four models. *Verify* is a sub method that can be further decomposed into four primitive steps including the perceptual primitive of *Look for Object in Viewing Space*, followed by the cognitive primitives of *Look for Object in Viewing Space*, *Compare Object to Perceived Memory*, and *Decision Step.*

The problem exists as subjects differed in where they placed the *verify* method. Most subjects and the baseline included *verify* as the last method for each model, while others included *verify* after each of the other sub methods in the model. An example is provided in table 5.

Table 4

Methods for Input Enemy Size

Baseline	Highlight Enemy Size Button	Select Description		Enter Quantity	Verify
Subject 1	Highlight Enemy Size Button	Select Description		Enter Quantity	Verify
Subject 2	Highlight Enemy Size Button	Select Description	Verify	Enter Quantity	Verify

Because verification can appropriately take place after each method or as the last method, these two models essentially represent two different methods for accomplishing the task. The purpose of this experiment is to compare models of the same method. For that reason, subject models that included multiple *verify* steps were modified to include only the final *verify*, thereby facilitating the comparison of the same method between subject models.

The steps were then compared to the baseline to determine accuracy and compared to other subject models to determine consistency. Additionally, the predicted time to accomplish each task was compared. Only the primitives defined in the baseline model were used for comparison; any other primitive defined by a subject that was not part of the baseline model was not included in the analysis.

4.8 Data Analysis

In this experiment, I am attempting to demonstrate the effectiveness of CAT-HCI as a tool capable of eliciting necessary information from subject matter experts while minimizing the involvement of the knowledge engineer throughout the process. I used three different analyses to determine the effectiveness of the CAT-HCI by comparing the accuracy of subject models, consistency between the various subject models, and accuracy of predicted times to accomplish each task. Appendix F contains the actual data reduction sheets used in calculating the various performance measure values.

4.8.1 Calculation Procedures for Accuracy of the Models

To measure accuracy, each model was compared with the baseline to determine the percent agreement similar to the example in Tables 3 and 4. Each of the 18 subjects created four models for a total of 72 models. One of the subject models was lost, thereby reducing the number of models for analysis to 71. The percent agreement as compared to the baseline for each of the 71 models was used to determine the mean level of agreement across all primitive tasks. Additionally, the differences in accuracy between physical, perceptual, and cognitive steps were compared. Table 5 provides an example of how accuracy measures were calculated using the sub method Select Activity for the Method Enemy Activity.

4.8.2 Calculation Procedures for Consistency Between Subject Models

Consistency describes how reliable subjects were in including or excluding a particular primitive step. All primitive steps found in the baseline model were considered for this measure. For every baseline primitive step of a method, the primitive steps generated by subjects were analyzed to determine presence or absence. Both the number of subjects who included the step and the number who did not include the step were counted. Consistency was determined by dividing the greater of the two values (number of subjects including or number of subject not including the step) by the total number of subjects. This process was completed for each of the 108 primitive steps across all four models. Table 6 provides an example of how consistency was measured using the sub method Highlight Activity for the method Enemy Activity.

4.8.3 Calculation Procedures for Consistency of Time Predictions

The predicted time for each operator was listed, with the sum of all the operators equal to the predicted time for the entire task. The predicted time was compared to the baseline time to determine percent agreement as compared to the baseline. Then all 71 percent agreements were used to calculate the mean. Table 8 provides an example of how predicted time measures were calculated using the sub method Select Activity for the method Enemy Activity.

4.8.4 Statistical Analysis

The results from all three measures, accuracy, consistency, and time were then compared to determine the mean (μ), the variance (σ^2), standard deviation (σ), and the confidence interval (CI). Additionally, a power analysis ($1-\beta$) of the experiment was determined to assess the reliability of the data. In this experiment we are interested in finding a minimal variance, a very tight confidence interval, and power equal to or in excess of 80%.

Table 5

Sample Calculations for Accuracy Measures

Method: Enemy Activity										
Select Activity										
Baseline	LFOVS		PWJ		PB		Select Activity Summary			
	Perc.	Cog.	Cog.	Motor	Cog.	Motor	Cog.	Perc.	Phys.	Total
Subject 1	LFOVS	LFOVS	PWJ	PWJ	PB	PB	100%	100%	100%	100%
Subject 2	LFOVS	LFOVS	PWJ	PWJ	PB	PB	100%	100%	100%	100%
Subject 3	LFOVS	LFOVS	PWJ	PWJ	PB	PB	100%	100%	100%	100%
Subject 4	LFOVS	LFOVS			PB	PB	67%	100%	50%	67%
Subject 5	LFOVS	LFOVS	PWJ	PWJ	PB	PB	100%	100%	100%	100%
Subject 6	LFOVS	LFOVS	PWJ	PWJ	PB	PB	100%	100%	100%	100%
Subject 7	LFOVS	LFOVS	PWJ	PWJ	PB	PB	100%	100%	100%	100%
Subject 8	LFOVS	LFOVS	PWJ	PWJ	PB	PB	100%	100%	100%	100%
Subject 9	LFOVS	LFOVS	PWJ	PWJ	PB	PB	100%	100%	100%	100%
Subject 10	LFOVS	LFOVS	PWJ	PWJ	PB	PB	100%	100%	100%	100%
Subject 11	LFOVS	LFOVS	PWJ	PWJ	PB	PB	100%	100%	100%	100%
Subject 12							0%	0%	0%	0%
Subject 13	LFOVS	LFOVS	PWJ	PWJ	PB	PB	100%	100%	100%	100%
Subject 14	LFOVS	LFOVS	PWJ	PWJ	PB	PB	100%	100%	100%	100%
Subject 15	LFOVS	LFOVS	PWJ	PWJ	PB	PB	100%	100%	100%	100%
Subject 16	LFOVS	LFOVS	PWJ	PWJ	PB	PB	100%	100%	100%	100%
Subject 17	LFOVS	LFOVS	PWJ	PWJ	PB	PB	100%	100%	100%	100%
Subject 18	LFOVS	LFOVS	PWJ*	PWJ*	PB	PB	100%	100%	100%	100%
					Mean		93%	94%	92%	93%

Legend	
Cognitive Operator	Physical Operator
Perceptual Operator	Combined

Table 6

Sample Calculations for Consistency Measures

Method: Input Enemy Activity							
	Highlight Activity						
Baseline	LFOVS		DS	PWJ		PB	
	Perc.	Cog	Cog	Cog.	Motor	Cog.	Motor
Subject 1							
Subject 2							
Subject 3				PWJ	PWJ	PB	PB
Subject 4				PWJ	PWJ		
Subject 5	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 6	DS	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 7	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 8	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 9	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 10	LFOVS	LFOVS	DS	PWJ	PWJ	PB	PB
Subject 11	LAVS	LAVS	DS	PWJ	PWJ	PB	PB
Subject 12	LFOVS	LFOVS				PB	PB
Subject 13			DS				
Subject 14	DS	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 15	LFOVS	LFOVS	PWJ	PWJ	PB	PB	DS
Subject 16	LFOVS	LFOVS	DS	PWJ	PWJ	PB	PB
Subject 17	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 18	LFOVS	LFOVS		PWJ*	PWJ*	PB	PB
# of subjects who included Step	11	11	4	13	13	13	13
# of subjects who did not include Step	7	7	14	5	5	5	5
# of subjects	18	18	18	18	18	18	18
% Consistency	61.1%	61.1%	77.8%	72.2%	72.2%	72.2%	72.2%

(light shading)	Cognitive	LFOVS	Look for Object in Viewing Space
(dark shading)	Perceptual	DS	Decision Step
(no shading)	Physical/Motor	PWJ	Point with Joystick

Table 7

Sample Calculations for Time Measures

Method: Input Enemy Activity Select Activity							
Baseline	LFOVS		PWJ		PB		Total Time
	Perc.	Cog.	Cog.	Motor	Cog.	Motor	
Subject 1	0.29	0.05	0.05	1.1	0.05	0.58	2.12
Subject 2	0.29	0.05	0.05	1.1	0.05	0.58	2.12
Subject 3	0.29	0.05	0.05	1.1	0.05	0.58	2.12
Subject 4	0.29	0.05			0.05	0.58	0.97
Subject 5	0.29	0.05	0.05	1.1	0.05	0.58	2.12
Subject 6	0.29	0.05	0.05	1.1	0.05	0.58	2.12
Subject 7	0.29	0.05	0.05	1.1	0.05	0.58	2.12
Subject 8	0.29	0.05	0.05	1.1	0.05	0.58	2.12
Subject 9	0.29	0.05	0.05	1.1	0.05	0.58	2.12
Subject 10	0.29	0.05	0.05	1.1	0.05	0.58	2.12
Subject 11	0.29	0.05	0.05	1.1	0.05	0.58	2.12
Subject 12							0
Subject 13	0.29	0.05	0.05	1.1	0.05	0.58	2.12
Subject 14	0.29	0.05	0.05	1.1	0.05	0.58	2.12
Subject 15	0.29	0.05	0.05	1.1	0.05	0.58	2.12
Subject 16	0.29	0.05	0.05	1.1	0.05	0.58	2.12
Subject 17	0.29	0.05	0.05	1.1	0.05	0.58	2.12
Subject 18	0.29	0.05	0.05	1.1	0.05	0.58	2.12
				Mean Predicted Time:			1.9383

Numbers are time in seconds

	Cognitive	LFOVS	Look for Object in Viewing Space
	Perceptual	PWJ	Point with Joystick
	Physical/Motor	PB	Push Button

5 ANALYSIS OF RESULTS AND INTERPRETATION

5.1 Analysis of Results

The cognitive models generated by the subjects were analyzed for accuracy and consistency. Accuracy measures were computed by comparing the primitive operators from each model with the primitive operators of the corresponding baseline model. Consistency measures were computed by comparing primitive operators from all models with the corresponding primitive steps generated by every other subject. Both of these analyses were conducted on the summary of total combined steps, as well as for each of primitive task subtype; cognitive, perceptual, and physical. The accuracy of time measures were computed by comparing the predicted time for each subject model with the predicted time for the corresponding baseline models.

5.1.1 Accuracy – Baseline Comparison

The results from the analysis of accuracy of the cognitive models are given in Table 8. The accuracy analysis was conducted on four sets of measures based on the type of primitive. They included cognitive, perceptual, and physical and combined. Cognitive primitives are the mental processes that occur throughout a task consisting of retrieval operations, mental activities, decisions etc. Examples of mental primitives are: *Compare Something to Perceived Memory*, and *Initiate Point with Joystick*. Perceptual Primitives consist of auditory, visual, or a tactile processes. Examples of perceptual operators include: *Look for Object in the Viewing Space*, *Look at Alphanumeric String* and *Listen to What You Are Hearing*. Physical primitives include eye-head movement, finger-hand-arm movement, and speech. Examples of physical primitives include: *Move Hand to Joystick* and *Press Button*. The combined measure includes the total primitive operators for a given model and includes cognitive, perceptual, and physical operators as appropriate.

The average number of total combined steps for a given task compared to the baseline was 76.9%. The average number of cognitive, perceptual and physical primitive steps defined by subjects for a given task as compared to the baseline model was 73.38% for cognitive, 78.64% for perceptual, and 83.98% for Physical.

5.1.2 Consistency – Comparison Between Subjects

The results from the analysis of consistency of the knowledge bases are given in Table 9. Similar to the measure of accuracy, the measure of consistency was analyzed across for the total combined steps as well as the three subsystems; cognitive, perceptual, and physical. The average level of agreement between subjects for all steps was 79.38%. The average level of agreement for the different subsystems was 76.52% for cognitive, 78.94% for perceptual, and 83.53% for physical.

5.1.3 Time – Baseline Comparison

The results from the analysis of accuracy of predicted time as compared to the baseline are given in Table 10. Results show the mean accuracy of time for each of the four models along with a summary of all 71 models. The average level of agreement between subject models and the baseline for predicted time to accomplish the tasks was 80.51%. The average time accuracy for the four models beginning with the first was 83.03%, 74.48%, 84.20%, and 80.45%.

Tables 11-13 show a summary of the respective confidence intervals for each measure.

Table 8

Summary of Accuracy based on 71 models

	Accuracy Summary of Results							
	Cognitive		Perceptual		Physical		Combined	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Subject 1	43.74%	23.21%	35.42%	30.71%	64.58%	29.17%	48.18%	24.53%
Subject 2	56.02%	18.15%	64.58%	21.92%	69.79%	15.73%	60.92%	18.16%
Subject 3	65.99%	14.90%	56.25%	26.68%	100.00%	0.00%	73.42%	13.60%
Subject 4	38.98%	25.91%	47.22%	31.55%	52.78%	29.27%	52.50%	28.47%
Subject 5	73.20%	10.86%	77.08%	15.77%	100.00%	0.00%	79.79%	9.77%
Subject 6	84.31%	13.76%	72.92%	7.98%	100.00%	0.00%	88.78%	5.08%
Subject 7	73.56%	12.00%	87.50%	15.96%	88.54%	7.89%	82.63%	11.79%
Subject 8	85.52%	17.04%	91.67%	16.67%	100.00%	0.00%	85.69%	13.39%
Subject 9	82.89%	9.48%	100.00%	0.00%	91.67%	9.62%	91.95%	4.74%
Subject 10	93.21%	9.43%	91.67%	16.67%	100.00%	0.00%	95.59%	8.82%
Subject 11	100.00%	0.00%	100.00%	0.00%	100.00%	0.00%	84.78%	30.43%
Subject 12	50.04%	14.99%	75.00%	31.91%	45.83%	14.43%	58.19%	12.14%
Subject 13	49.79%	16.75%	43.75%	18.48%	47.92%	17.18%	51.54%	20.46%
Subject 14	83.48%	19.42%	93.75%	12.50%	79.17%	25.00%	89.09%	19.03%
Subject 15	95.52%	3.18%	100.00%	0.00%	94.79%	6.25%	96.96%	3.61%
Subject 16	100.00%	0.00%	100.00%	0.00%	100.00%	0.00%	97.83%	4.35%
Subject 17	80.83%	7.57%	100.00%	0.00%	100.00%	0.00%	81.25%	14.23%
Subject 18	55.14%	9.74%	70.83%	20.97%	68.75%	27.53%	60.72%	14.62%
Mean	72.90%		78.20%		83.55%		76.66%	
Std. Dev.	19.80%		21.43%		20.01%		16.89%	
90% CI	8.12%		8.79%		8.21%		6.93%	

	All 71 Models			
Mean	73.38%	78.64%	83.98%	76.88%
Std. Dev	22.68%	26.15%	23.15%	21.86%
90% CI	4.49%	5.18%	4.58%	4.33%

Table 9

Summary of Consistency Between 18 Subjects on 108 Steps

		Consistency			
		Cognitive	Perceptual	Physical	Total Combined
Model 1 – Size (23 Steps)	Mean	76.05%	78.43%	84.31%	78.52%
	Std. Dev.	15.90%	17.97%	9.61%	14.58%
Model 2 – Activity (17 Steps)	Mean	75.00%	72.22%	83.33%	76.47%
	Std. Dev.	15.55%	19.25%	12.83%	15.15%
Model 3 – Location (40 Steps)	Mean	79.29%	81.48%	82.87%	80.69%
	Std. Dev.	12.96%	12.99%	10.97%	12.20%
Model 4 - Unit (28 Steps)	Mean	77.78%	80.56%	84.03%	79.96%
	Std. Dev.	13.76%	14.70%	6.26%	12.13%
Summary of all Models (108 Steps)	Mean	76.52%	78.94%	83.53%	79.38%
	Std. Dev.	14.03%	14.29%	9.40%	13.10%
	90% CI	2.98%	6.26%	2.87%	2.09%

Table 10

Summary of Predicted Times for 71 Models

	Time % Agreement w/ Baseline				
	Size	Activity	Location	Unit	Total
Baseline Time (sec)	7.15	5.73	13.605	9.435	35.92
Subject 1	77.62%	41.62%	30.06%	63.01%	53.08%
Subject 2	69.09%	41.62%	79.13%	83.25%	68.27%
Subject 3	77.62%	72.08%	90.74%	86.65%	81.77%
Subject 4	No Data	41.62%	81.04%	28.46%	50.37%
Subject 5	82.38%	78.01%	93.24%	86.65%	85.07%
Subject 6	82.38%	94.07%	97.50%	86.65%	90.15%
Subject 7	95.52%	78.01%	81.04%	96.61%	87.79%
Subject 8	100.00%	78.01%	100.00%	100.00%	94.50%
Subject 9	82.66%	86.91%	81.66%	92.05%	85.82%
Subject 10	100.00%	80.98%	100.00%	100.00%	95.24%
Subject 11	100.00%	100.00%	100.00%	100.00%	100.00%
Subject 12	44.41%	44.15%	70.56%	59.62%	54.68%
Subject 13	62.03%	41.62%	72.69%	38.21%	53.64%
Subject 14	81.54%	100.00%	100.00%	63.91%	86.36%
Subject 15	100.00%	100.00%	97.65%	96.61%	98.56%
Subject 16	100.00%	100.00%	100.00%	100.00%	100.00%
Subject 17	87.13%	83.94%	93.24%	90.25%	88.64%
Subject 18	69.09%	78.01%	47.04%	76.26%	67.60%
Mean	83.03%	74.48%	84.20%	80.45%	80.09%
Variance	2.50%	5.04%	3.82%	4.66%	3.03%
Std/ Dev.	15.80%	22.45%	19.55%	21.59%	17.42%
90% CI	6.67%	9.21%	8.02%	8.85%	7.14%

All 71 Models	
Mean	80.51%
Standard Deviation	20.00%
90% Confidence Interval	3.96%

Table 11

Confidence Intervals for Accuracy

	Accuracy			
	% Confidence	CI	Lower Limit	Upper Limit
Cognitive	90%	8.12%	64.78%	81.02%
	95%	9.85%	63.05%	82.75%
	99%	13.52%	59.38%	86.42%
Perceptual	90%	8.79%	69.41%	86.99%
	95%	10.66%	67.55%	88.86%
	99%	14.64%	63.57%	92.84%
Physical	90%	8.21%	75.34%	91.75%
	95%	9.95%	73.59%	93.50%
	99%	13.67%	69.88%	97.22%
Combined	90%	7.81%	68.61%	84.23%
	95%	9.47%	66.95%	85.90%
	99%	13.01%	63.41%	89.43%

Table 12

Confidence Intervals for Consistency

	Consistency			
	% Confidence	CI	Lower Limit	Upper Limit
Cognitive	90%	2.98%	74.05%	80.01%
	95%	3.56%	73.47%	80.59%
	99%	4.74%	72.29%	81.77%
Perceptual	90%	6.26%	71.91%	84.44%
	95%	7.57%	70.60%	85.75%
	99%	10.53%	67.64%	88.70%
Physical	90%	2.92%	80.72%	86.55%
	95%	3.51%	80.13%	87.15%
	99%	4.73%	78.90%	88.37%
Combined	90%	2.09%	76.82%	81.00%
	95%	2.50%	76.41%	81.41%
	99%	3.30%	75.61%	82.21%

Table 13

Confidence Intervals for Predicted Time

	Time			
	% Confidence	CI	Lower Limit	Upper Limit
	90%	6.48%	76.55%	89.51%
Model 1	95%	7.86%	75.17%	90.89%
	99%	10.79%	72.23%	93.82%
	90%	9.21%	65.27%	83.69%
Model 2	95%	11.17%	63.32%	85.65%
	99%	15.34%	59.15%	89.82%
	90%	8.02%	76.18%	92.21%
Model 3	95%	9.72%	74.48%	93.92%
	99%	13.35%	70.85%	97.55%
	90%	8.85%	71.60%	89.31%
Model 4	95%	10.74%	69.72%	91.19%
	99%	14.74%	65.71%	95.20%
	90%	3.96%	76.55%	84.46%
Total	95%	4.75%	75.76%	85.25%
	99%	6.31%	74.19%	86.82%

A level of 80% agreement and accuracy was considered desirable at the beginning of this experiment. In addition to calculating the confidence intervals as described in the previous tables, the data was further compared with the goal of 80% to determine the likelihood of falsely rejecting the H_0 ($\mu < 80\%$) and/or falsely accepting the H_0. Using an α of .05 across all the measures, the cognitive accuracy ($\mu = 73.38\%$) was the only instance where the H_0 was rejected. Time measures, cognitive consistency, and all perceptual, physical and combined measurements regarding accuracy and consistency were statistically equal to or greater than 80% and therefore could not be rejected. Finally, a power analysis was conducted to determine the possibility of falsely accepting the H_0. The βs are listed in Table 13.

Table 14

Summary of β

Accuracy β		Consistency β		Time β	
Cognitive	20.3%	Cognitive	49.0%	Task 1	0.7%
Perceptual	11.5%	Perceptual	4.82%	Task 2	27.4%
Physical	*	Physical	*	Task 3	0.5%
				Task 4	4.2%
Total	38.5%	Total	7.42%	Total	17.5%

* Both physical tasks were statistically greater than 80%, and therefore would not falsely be accepted.

5.2 Interpretation of Results

Subjects successfully used CAT-HCI to create models that were nearly 80% accurate with regard to all steps combined, as well as for physical and perceptual operators. This overall accuracy of between 75%-80% represents a potential savings of three-quarters of a knowledge engineer's time, resulting in less costly knowledge base development and more affordable adaptive training systems.

The consistency between models was also very high at 79.4%. Consistency shows that the subject models were on average 79.4% similar. In other words, their models contained 79.4% of the same primitive operators for any given method. Consistency was calculated to determine the ability to generate common knowledge bases from multiple domain experts where no baseline model exists. Based on the consistency data, one can infer that accurate cognitive models can be created for domain tasks for which baseline models do not exist. Differences in the level of detail between models would then be clarified during the verification stage. The five stages of knowledge acquisition are: 1) elicitation, 2) organization, 3) representation, 4) refinement, and 5) verification. (Kotnour, 1992). This experiment focused on the tool's ability to elicit the information from domain experts. Based on subject input, the tool organized and represented the elicited information. The subjects were then able to edit and refine their models upon completion. In general, domain experts lack the knowledge acquisition skills to adequately verify their own models. Consequently, a trained analyst is required to verify the model by executing it and finding any inaccuracies in the model.

The average predicted time measures were 80.51% accurate when compared to the baseline models. This difference would most likely be corrected during the verification phase as the missing operators are identified. Unfortunately, no data exist on the actual time to conduct these methods making it impossible to verify the predicted time for each method. An interesting follow-on experiment would be to determine the actual time to conduct each task, and thereby provide a basis of comparison between predicted and actual times to determine its level accuracy.

5.2.1 Explicitly and Implicitly Derived Operators

Out of the possible 108 steps, 50% or more of the subjects included all but ten steps. Of these ten operators that were not included by a majority of the subjects, nine were cognitive in nature and one was perceptual. This is consistent with both the accuracy and consistency measures that found the cognitive subsystem to generally be the lowest measure. Of the 62 cognitive steps from all four models, 49 were implicitly generated by CAT-HCI based on input for a higher-level task. For example, when a subject accurately selected *push button*, CAT-HCI entered both the mental operator to initiate the button push as well as the physical action of pushing the button for a total two operators. The remaining 13 cognitive operators required explicit input from the subjects and included *decision step*, and *compare perceived object to memory*. Of the nine cognitive operators missed by a majority of subjects, eight required explicit input from the subject. Clearly CAT-HCI has made significant improvements in being able to implicitly include cognitive steps. Of the 13 explicitly required cognitive steps, the majority of subjects failed to identify 8. In other words, a majority of subjects only identified 38.5% of those cognitive steps that required explicit input by the subject. This further corroborates the difficulty that subject matter experts, not skilled in knowledge acquisition, have in explicitly identifying cognitive steps. An analysis of only these 13 cognitive steps that require explicit input by the subjects shows an accuracy of 40.93% and a consistency of 60.38%. Clearly, the majority of subjects lacked the knowledge engineering background to identify more than 45%

of the cognitive primitive operators. Since the actual level of cognitive accuracy was significantly higher at 73.38%, it can be concluded that CAT-HCI was responsible for eliciting a much higher level of accuracy for cognitive tasks than possible by domain experts without the aid of the tool.

5.2.2 A Look at Cognitive Operators

Subject matter exerts, with little understanding of knowledge engineering techniques, used CAT-HCI to create cognitive models that were on average 76.88% accurate when compared with the baseline. The four models had between 17 and 40 steps each with a total of 108 primitive steps. As mentioned above, more than half of the steps (62) were cognitive in nature, which proved to be the most difficult to capture based on the previous CAT evaluation where only 28.3% of the cognitive primitive tasks were identified. (Kotnour, 1992) CAT-HCI demonstrated a significant improvement by accurately eliciting 73.38% of the cognitive primitives. The results of the two experiments cannot be directly compared as the pool of subjects and the tasks were significantly different. However, it is useful to observe the results of both experiments, in particular as they relate to cognitive primitive tasks. CAT-HCI appears to have significantly improved the ability to infer information from subjects concerning cognitive primitives. Cognitive primitives are non-behavioral steps taken to accomplish a task and are implicitly performed by subjects, i.e. initiate button push. Since the test subjects were not cognitive analysts, it is unlikely that they would have defined many of these cognitive primitives on their own, yet CAT-HCI was able to elicit nearly 75% of those primitive from the domain experts. Mental primitives are particularly important in predicting the time to execute a cognitive task. The results of this experiment clearly indicate that CAT-HCI is able to elicit the necessary information from domain experts to create relatively accurate models.

6 CONCLUSIONS AND RECOMMENDATIONS

6.1 Significance of Research

This research provides the only evaluation of an automated cognitive analysis tool to determine accuracy and consistency of created models. As technology increases, there is an expectation for better and more adaptive training devices. While the technology exists to provide the necessary fidelity for many of these systems, AT systems must rely on cognitive models based on domain expertise. Unfortunately the high cost and inefficiency of developing these models have made AT systems unaffordable to most organizations. An automated tool for developing cognitive models must be evaluated and validated before it can be used to develop models for an AT system. This research provides the first evaluation of such a tool, thereby demonstrating its ability to create knowledge bases with an accuracy of between 75 and 80%.

6.2 Areas for Further Research

Recommendations for future research in this area are as follows:

1. Validate time predictions of CAT-HCI. CAT-HCI provides predicted times to accomplish particular methods based on the sum of times to accomplish a method's primitive operators. These times are based on various studies and are expected to be highly accurate. However, it would be beneficial to validate these times by conducting an experiment that compares the actual time a subject takes to accomplish a task with the times predicted by CAT-HCI generated models.

2. Determine the cause for the significant level of variance between subject models and also determine the level of training required to raise accuracy of models to a level of 90% or higher.

3. Conduct a Human Computer Interface Study to determine the most effective techniques for eliciting information from subject matter experts.

4. Develop an analogy system that would allow the tool to better derive information from the subject by comparing previously completed knowledge bases. This would potentially allow the tool to adapt or update the current model to previous models, and reduce the repetition of defining recurring methods that occur in multiple models.

6.3 Lessons Learned

I would make the following changes if I were to run this experiment again:

1. Provide subjects with a print out of their models for review before they finish making their models. This would also provide a point of reference and help them better understand the overall model.

2. Collect more detailed biographical data of subjects, specifically in relation to experience on computers to help identify differences in performance. There was a significant difference between subject performances. Some subjects maintained 100% accuracy for all models, while others had averages below 50%. Based on conversations with subjects, it appeared that those who were more familiar and comfortable with computers created more accurate models. More detailed biographical data might provide more accurate insight into this area.

3. Take better notes on subject comments and observations of subjects while conducting the task analysis. This would provide a better basis for potential improvements to the tool and might help explain differences in subject performance.

6.4 Summary of Major Outputs

The following is a list of the outputs of conducting this research.

1. A complete set of models for the Bradley A3 SPOT Report thumb cursor method.

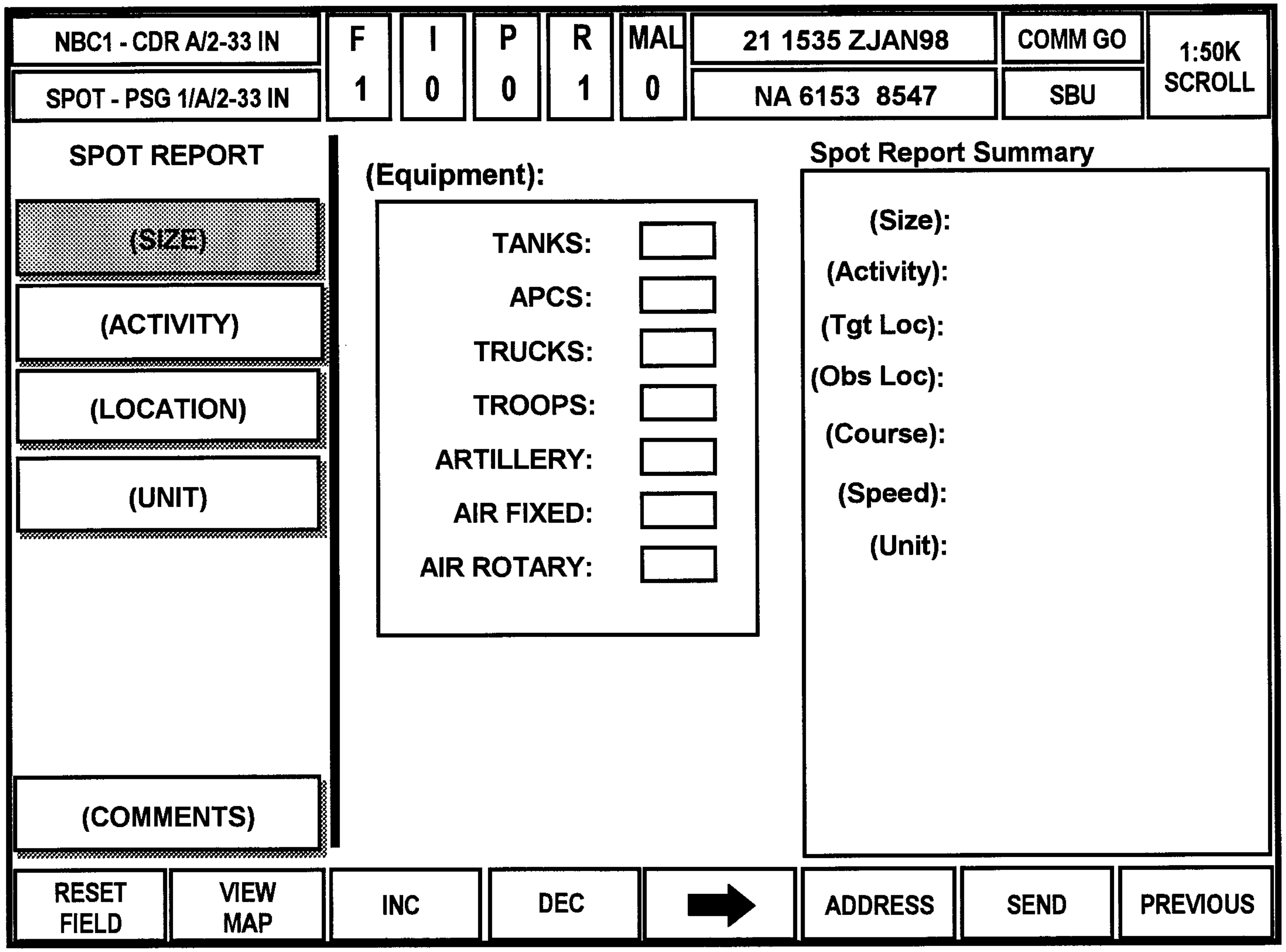
2. An evaluation of the CAT-HCI automated knowledge acquisition tool in terms of accuracy of subject models, accuracy of predicted time to execute a task, and consistency of models between subject matter experts generated by using the tool.
3. An evaluation methodology for evaluating knowledge bases.

APPENDIX A-LIST OF CTD TASKS

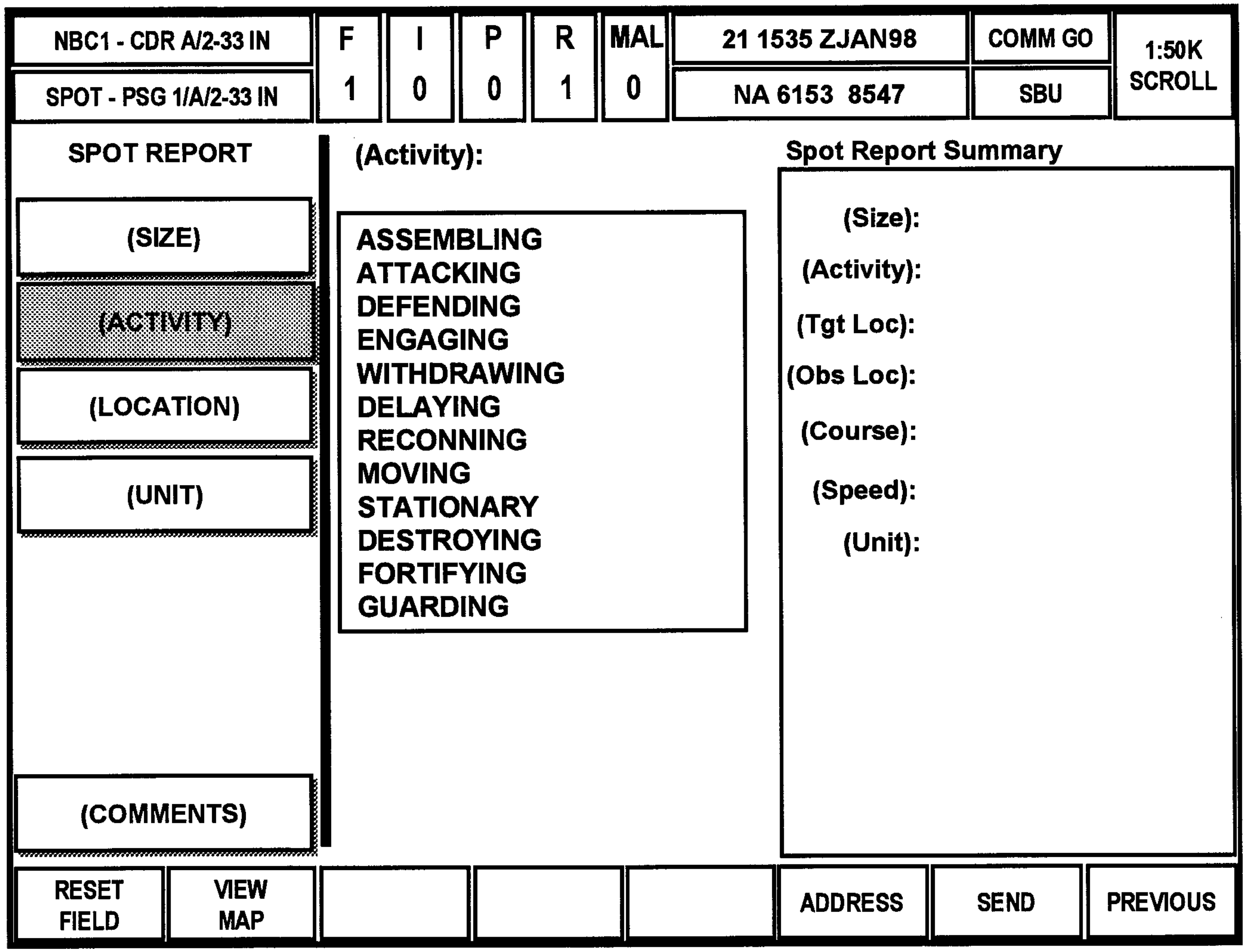
LIST OF CTD TASKS

Maintenance and Logistics	***"Digitization"/Command and Control***
Power Management	Read Message
Environmental Parameters	Create a Route
Manual Boresight	Free Text Message
Reticle Selection	Position Report
Fire Control MPL	Receive and Alert Message
Log-in	Send an Alert Message
Change Duty Position	Create and OPORD
Address Setup	Read and OPORD
Auto Reports	SPOT Report
Command & Control MPL	Situation Report
Map Setup	NBC 1 Nuclear
Navigation Setup	NBC 1 Chemical
Zeroize	NBC 1 Biological
Silent Watch	
Modes of Operation	
Primary System Setup	

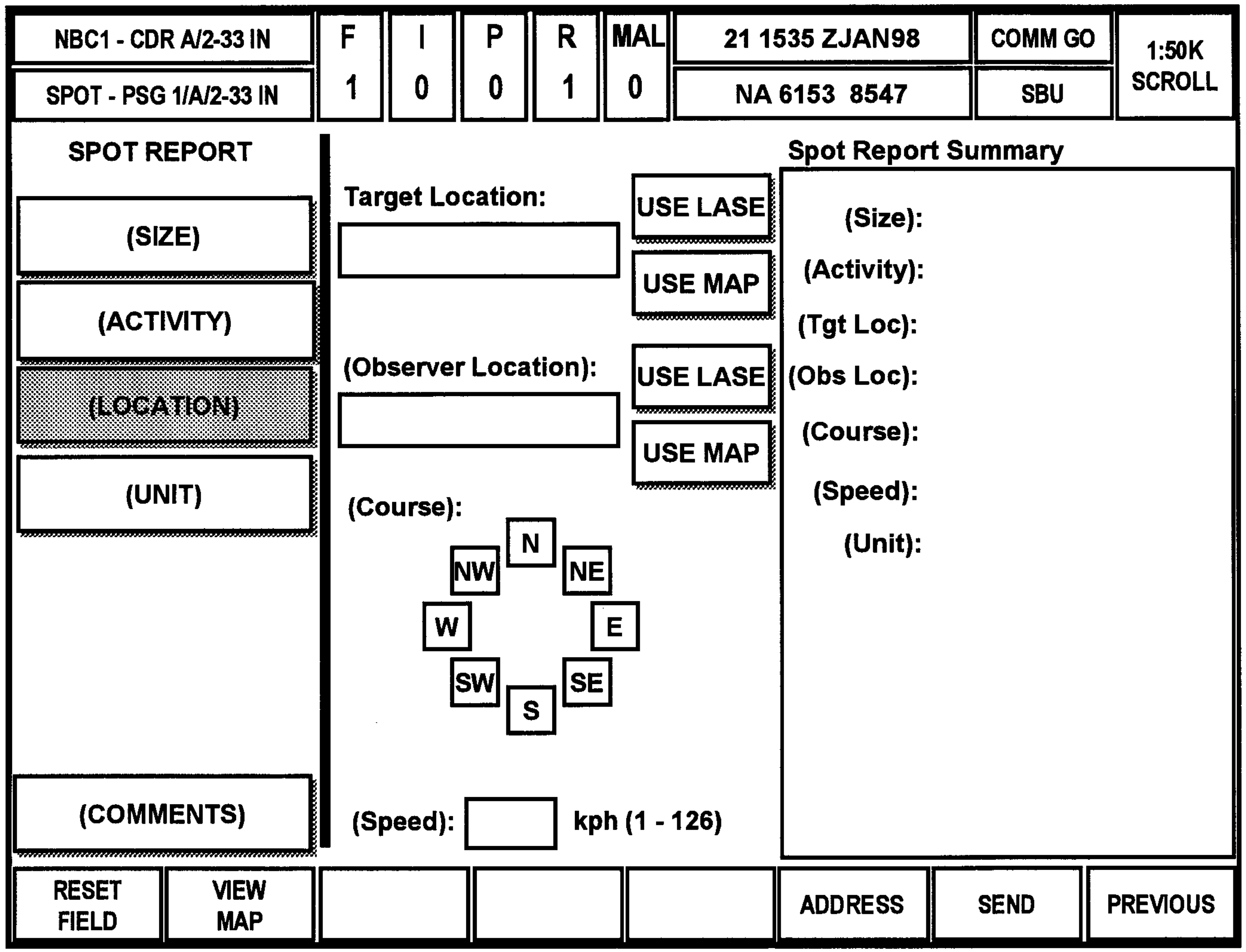
APPENDIX B – SCREEN TEMLATES AND BASELINE MODEL FOR CREATE SPOT REPORT



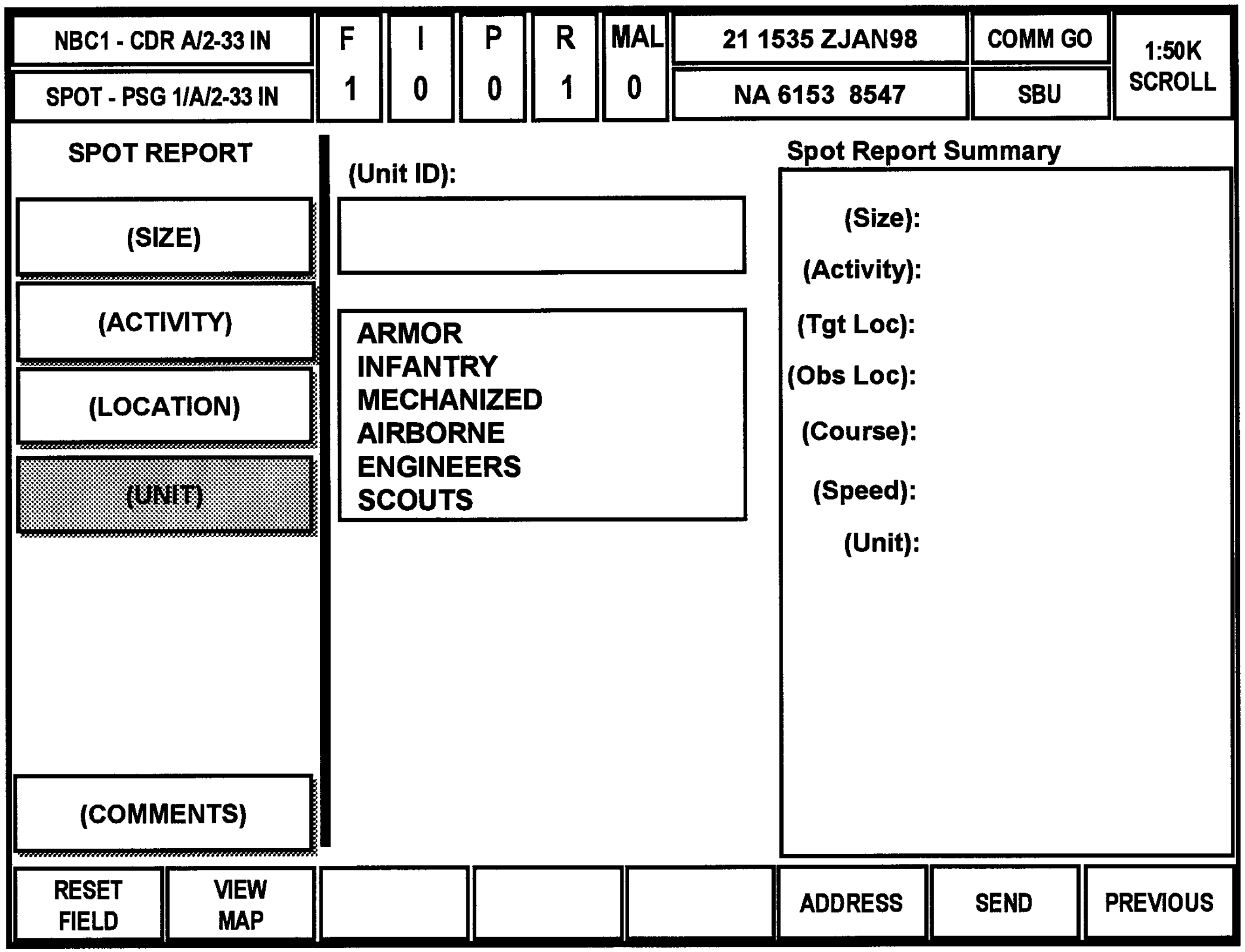
SPOT Report: Enemy Size



SPOT Report: Enemy Activity



SPOT Report: Location



SPOT Report: Enemy Unit

BASELINE MODEL FOR CREATE SPOT REPORT

1. Method - Input Enemy Size and Description

1.1. Verify SIZE is selected (highlighted)
  1.1.1. Look for Object in Viewing Space
  1.1.2. If SIZE is selected then go to Select Description (13), else continue (1.2)
1.2. Highlight SIZE
  1.2.1. Look for Object in viewing Space
  1.2.2. Point with the joystick
  1.2.3. Push Button
1.3. Select Description
  1.3.1. Look for Object in viewing Space
  1.3.2. Point with the joystick
  1.3.3. Push Button
1.4. Enter Quantity
  1.4.1. Move hand to keyboard (CTED)
  1.4.2. Type Natural Language Text (Input quantity)
1.5. If more enemy equipment go to Select Description (2.3), else continue (2.6)
1.6. Verify Summary on Right Side of Screen
  1.6.1. Look for Object in viewing Space
  1.6.2. Compare Perceived Object to Memory
  1.6.3. If incorrect go to Highlight Size

2. Method - Input Enemy Activity

2.1. Verify ACTIVITY is selected (highlighted)
  2.1.1. Look for Object in Viewing Space
  2.1.2. If ACTIVITY is selected then go to input Enemy Activity (2.3), else continue (2.2)
2.2. Highlight ACTIVITY
  2.2.1. Look for Object in viewing Space
  2.2.2. Point with the joystick
  2.2.3. Push Button
2.3. Input Enemy Activity
  2.3.1. Look for Object in viewing Space
  2.3.2. Point with the joystick
  2.3.3. Push Button
2.4. Verify Summary on Right Side of Screen
  2.4.1. Look for Object in viewing Space
  2.4.2. Compare Perceived Object to Memory
  2.4.3. If incorrect go to Input Enemy Activity

3. Method - Input Enemy Location

3.1. Verify LOCATION is selected (highlighted)
  3.1.1. Look for Object in Viewing Space
  3.1.2. If LOCATION is selected then go to Enter Target Location (3.3), else continue (3.2)
3.2. Highlight LOCATION
  3.2.1. Look for Object in viewing Space
  3.2.2. Point with the joystick
  3.2.3. Push Button
3.3. Enter Target Location
  3.3.1. If Conditions are "Accuracy is most important" or "speed is most important" then use method "Use Lase" (4.2.3)
  3.3.2. If Conditions are "Stealth is most important" or "Enemy cannot be lased" then use Method "Map" (4.2.4)
  3.3.3. Method – Use Lase
    3.3.3.1. Look for Object in viewing Space
    3.3.3.2. Point with the joystick
    3.3.3.3. Push Button
  3.3.4. Method – Use Map
    3.3.4.1. Look for Object in viewing Space
    3.3.4.2. Point with the joystick
    3.3.4.3. Push Button
  3.3.5. Method – Manual
    3.3.5.1. Look for Object in viewing Space
    3.3.5.2. Point with the joystick
    3.3.5.3. Push Button
    3.3.5.4. Move hand to keyboard
    3.3.5.5. Type Natural Language
3.4. Enter Observer's Location
  3.4.1. If Conditions are "Observer's Position Can be lased" and "Lasing the observer will not compromise position" and The Bradley Eysesafe Laser is used" then use method "Use Lase" (4.3.3)
  3.4.2. If Conditions are "Cannot Lase observer" then use Method "Map" (4.3.4)
  3.4.3. Method – Use Lase
    3.4.3.1. Look for Object in viewing Space
    3.4.3.2. Point with the joystick
    3.4.3.3. Push Button
  3.4.4. Method – Use Map
    3.4.4.1. Look for Object in viewing Space
    3.4.4.2. Point with the joystick
    3.4.4.3. Push Button
  3.4.5. Method – Manual

3.4.5.1. Look for Object in viewing Space
3.4.5.2. Point with the joystick
3.4.5.3. Push Button
3.4.5.4. Move hand to keyboard
3.4.5.5. Type Natural Language

3.5. Enter Course
3.5.1. If Enemy is moving then enter Course and Speed (Continue to 4.4.2), else Go To "Verify Summary" (4.6)
3.5.2. Look for Object in viewing Space
3.5.3. Point with the joystick
3.5.4. Push Button

3.6. Enter Speed
3.6.1. Look for Object in viewing Space
3.6.2. Point with the joystick
3.6.3. Push Button
3.6.4. Type Speed
3.6.4.1. Move Hand to Keyboard (CTED)
3.6.4.2. Type Natural Language Text (Input Speed)

3.7. Verify Summary on Right Side of Screen
3.7.1. Look for Object in viewing Space
3.7.2. Compare Perceived Object to Memory
3.7.3. If incorrect go to Enter Target Location (3.3)

4. Method - Input Enemy Unit Description

4.1. Verify UNIT is selected (highlighted)
4.1.1. Look for Object in Viewing Space
4.1.2. If UNIT is selected then go to Enter Unit ID (4.3), else continue (4.2)

4.2. Highlight UNIT
4.2.1. Look for Object in viewing Space
4.2.2. Point with the joystick
4.2.3. Push Button

4.3. Enter Unit ID
4.3.1. If Precise Unit Description is known, then Enter Unit ID (5.2.2), else Go To "Select Unit Description (5.2.3)
4.3.2. Enter Unit ID
4.3.2.1. Look for Object in viewing Space
4.3.2.2. Point with the joystick
4.3.2.3. Push Button
4.3.2.4. Move hand to keyboard (CTED)
4.3.2.5. Type Natural Language Text (Input ID)
4.3.3. Select Unit Description
4.3.3.1. Look for Object in viewing Space
4.3.3.2. Point with the joystick

4.3.3.3. Push Button

4.4. Verify Summary on Right Side of Screen

4.4.1. Look for Object in viewing Space

4.4.2. Compare Perceived Object to Memory

4.4.3. If incorrect go to Enter Unit ID (4.3)

APPENDIX C – PROTOCAL FOR SCREENING TASK SESSION

PROTOCOL FOR SCREENING TASK
SESSION

Using the Bradley Desktop Trainer (BDT), up to four subjects can be screened at a time.

All four can all be observed on the Instructor/Operator (I/O) Station.

1. Initialize Bradley Desktop Trainer (BDT) with Main Menu Showing.
2. Have the subject fill out the questionnaire.
3. Explain purpose of the screening session to the subject.

   The purpose of this session is to have you perform a task that you will later be asked how to perform. This is done to ensure that you can do the task.

   You will use the BDT to perform the task as defined by the piece of paper I will give you. I will observe you from the I/O station while you are performing the task. If you do not properly perform the task the first time, you will be excused to refresh/retrain yourself on the task in order to be retested again with a subsequent group. If you fail the task a second time, you will be removed from the experiment.

4. Give subject written explanation of the task.
5. Have the subject perform the task.
6. Observe subject perform the task from the Instructor/Operator (I/O) Station.
7. Check the subject's SPOT Report Summary against the solution to ensure he performed the task correctly, then begin next session. If the subject did not perform the task correctly, then send him out to be retested later. If a subject fails the test a second time, he will be eliminated from the experiment.

Cognitive Analysis Tool
Subject Information Questionnaire

Subject #___

Name: ______________________________ Rank: _______________

Organization: _________________________ Phone: _______________ Email: ____________

1. Military Status?

   Active Duty | Retired | Former Military | No Military Service

2. Current Employment?

   US Army | United Defense | Other _______________

3. Infantry/Armor Experience?

   N/A | 0-5 Years | 6-14 Years | 15 or more Years

4. How long have you been associated with the Bradley A3?

   Less than 3 Months | 4-12 Months | 1-2 Years | More than 2 years

5. How are you involved with the Commander's Tactical Display (CTD)? Mark all that apply.

   Instructor | Screen Development | Other_______________

6. How frequently do you interact with the CTD?

   Every Day | Weekly | Monthly | Less than once a month

Task Explanation for Screening Session

During this session, you are to create a SPOT Report based on the following observation. Use the Map method to pluck a grid designate enemy location. Stop when you have entered all enemy information. ***Do not send the report.***

Observed Enemy Activity

You are observing three (3) T-72 tanks and four (4) BMPs moving northeast at a speed of approximately 20kph. Their current location is (last lased position).

(Alternate) Task Explanation for Screening

Session

During this session, you are to create a SPOT Report based on the following observation. Use the Map method to pluck a grid designate enemy location. Stop when you have entered all enemy information. ***Do not send the report.***

Observed Enemy Activity

You are observing four (4) trucks and ten (10) troops withdrawing east at a speed of approximately 5kph. Their current location is (last lased position).

Score Sheet for Subject Screening Session

Subject #____

Review completed SPOT Report and verify the following information in the summary located on the right side of the screen.

______1.	(SIZE):	3 Tanks and 4 APCS
______2.	(ACTIVITY):	MOVING
______3.	(Tgt Location):	+ or – 200m of actual location
______4.	(Obs Location):	+ or – 100m of actual location (Own Location)
______5.	(Course):	NE
______6.	(Speed):	20kph
______7.	(Unit):	Mechanized, Scouts, or Blank

APPENDIX D – PROTOCAL FOR CAT-HCI FAMILIARIZATION

EXPLANATION FOR CAT-HCI
FAMILIARIZATION

1. Explain the purpose of this Session. The purpose of this session is to familiarize you with the CAT-HCI (Cognitive Analysis Tool - Human Computer interaction) Tool. The session will consist of three parts. 1) I will define the top-level terms and processes used by CAT-HCI. 2) We will walk through an example together. 3) You will each create an example cognitive model of your own.

2. Define Key Terms used in CAT-HCI.

    a. Top Level Goal – The main task we are trying to describe. CAT-HCI will elicit the knowledge about this top-level goal so that all steps are listed. An example of a top level goal "Use the CTD to Power-up the A3 System."

    b. Goal – An action or task that is described by steps. An example is "Use the CTD to Power-up the A3 System."

    c. Steps – Action or task taken to accomplish a goal. Examples are "Log-in the system," and "Enter Date Time Group."

    d. Method – A set of steps to accomplish a particular goal. Used to describe or group the steps. In some cases, more than one method may be used to accomplish a particular goal.

    e. Selection Rule – Specifies the conditions under which a specific method is used to accomplish a goal when alternative methods exist. An example is: "The Date Time Group is incorrectly displayed."

    f. Primitive – A step that cannot be decomposed into further substeps. Primitives can be either mental or physical. Examples of mental primitive are "Decision between alternatives," and "Compare something to perceived memory." Examples of physical primitive are "point, using joystick," and "press button."

3. Do a walk through example (NBC 1 Report – Access & NBC Select)

    a. Create a new model using Guidance Mode. Tutorial Mode does not work.
        1) Press Guidance Mode
        2) Select New

    b. Top Level Goal Screen
        1) Explain help files not available.
        2) Purpose of this screen is to name the top-level goal.
        3) Use the Goal Name that has been provided. (Description is optional)

    c. Method Editor
        1) Purpose: to define the steps to accomplish the top-level goal.
        2) Name of Current Step to accomplish goal

3) Description (optional)
4) Move Step to List
5) Explain Edit, Delete, Cancel Buttons
6) Explain only one action per step
7) Press Done button when all steps have been listed.
8) Do Example Steps for Access NBC1 Report

d. Order the Steps
   1) Purpose – define the order of the steps
   2) As entered
   3) No specific Order
   4) Another Order (Reorder Method's Step Information)
      a) Add (moves step to reordered side, bottom of list),
      b) Remove (removes step from reordered side)
      c) Insert (inserts above highlighted step on the reordered side)
      d) Clear (Clears reordered side)
      e) Done

e. Define Alternative Methods
   1) Purpose – Define another set of steps
   2) Yes
      a) Requests names for current and previous method
      b) Results in method editor to define a new set of steps
   3) No – proceeds to next method
   4) Do Example (Lase vs. Map Method for selecting target location)

f. Definition of Selection Rule
   1) Purpose – Define criteria for selecting each method
   2) "And" vs. "Or" - Enter And conditions in same dialog box. Enter Or conditions individually as "Alternative Selection Rules."
   3) Do example (Conditions for Lase and Conditions for Map)

g. Define Primitive Methods
   1) Use Pull down menus from Method editor to define Primitives
      a) Insert HCI Operator
         (1) Arm/Hand/Finger
         (2) Visual
         (3) Auditory
         (4) Motor Speech
         (5) Mental
         (6) Wait for System
      b) Insert Standard Operator
         (1) Subgoal
         (2) Decision Step
         (3) Go To Step

2) Do Example

h. Complete Example Problem

4. Subjects complete NBC1 Attack on by themselves. Provide feedback to the subject.

APPENDIX E- PROTOCAL FOR MACHINE AIDED SESSSION

Task Explanation for Machine Aided Session

During this session you will use the tool called CAT-HCI. The tool will prompt you for information about how to accomplish five methods to accomplish a goal.

Remember to be as precise as possible. The CAT-HCI tool should prompt you to include the necessary primitives to create an effective model. **Create a new model for each method.**

The Goal is: Create a SPOT Report from the Bradley A3, CTD.

The Methods are:

1. Enter Enemy Size and Description
2. Enter Enemy Activity
3. Enter Enemy Location
4. Enter Enemy Unit Description

You can begin this session by double-clicking the CAT-HCI icon and selecting the **Guidance** mode and selecting a new model with top-level goal of "Access SPOT Report." Upon completion of each method, save it and then create a new file for the next method until you have completed all five methods. Save each method as a file with the following names, where XXX refers to your initials:

1. SizeXXX.hci
2. ActXXX.hci
3. LocXXX.hci
4. UnitXXX.hci

Assumptions:

1. The Bradley A3 is powered on.
2. The CTD is on and at the main menu.
3. You have already logged onto the system.
4. For standardization, the following primitives should be used for this session:

	Action	Use CAT HCI Primitive
Physical Primitive	Press Hard Keys	Press button
	Point, using thumb cursor	Point, using joy stick
	Press thumb cursor	Press button
Mental Primitives	Find Specific Icon/object	Look for object in viewing space

Table 7. Defined BDT Primitives

Example use of the primitives:

Goal: Select the Direction of Attack from the NBC Selection Screen

step 1. Look for object in viewing space

step 2. Point using joy stick (thumb cursor control)
step 3. Press button (click thumb cursor to select item)

APPENDIX F – DATA REDUCTION SHEETS

Model 1	Method: Highlight Size						
Baseline	LFOVS		DS	PWJ		PB	
	Perc.	Cog.	Cog.	Cog.	Motor	Cog.	Motor
Subject 1	LFOVS	LFOVS	DS	PWJ*	PWJ*	PB	PB
Subject 2	LFOVS	LFOVS	DS	PWJ	PWJ	PB	PB
Subject 3	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 5	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 6	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 7	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 8	LFOVS	LFOVS	DS	PWJ	PWJ	PB	PB
Subject 9	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 10	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 11	LFOVS*	LFOVS*	DS	PWJ	PWJ	PB	PB
Subject 12							
Subject 13			DS				
Subject 14	DS	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 15	LFOVS	LFOVS	DS	PWJ	PWJ	PB	PB
Subject 16	LFOVS	LFOVS	DS	PWJ	PWJ	PB	PB
Subject 17	LFOVS	LFOVS	PWJ	PWJ	PB	PB	DS
Subject 18	LFOVS	LFOVS		PWJ	PWJ	PB	PB

* Subject used semantically equivalent primitive operator

CPOTM	Compare Object to Memory
DS	Decision Step
LFOVS	Look for Object in Viewing Space
MHTKB	Move Hand to Keyboard
PB	Press Button
PWJ	Point with Joystick
RWS	Read What you See
TC	Type Code
TNL	Type Natural Language

Subsystem
- Cognitive
- Perceptual
- Physical
- Combined

Model 1	Method: Select Description					
Baseline	LFOVS		PWJ		PB	
	Perc.	Cog.	Cog.	Motor	Cog.	Motor
Subject 1			PWJ	PWJ	PB	PB
Subject 2	LFOVS	LFOVS	PWJ	PWJ		
Subject 3			PWJ	PWJ	PB	PB
Subject 5	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 6	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 7	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 8	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 9	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 10	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 11	LFOVS*	LVOVS*	PWJ	PWJ	PB	PB
Subject 12	LFOVS	LFOVS	PWJ	PWJ		
Subject 13	LFOVS	LFOVS	PWJ	PWJ		
Subject 14	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 15	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 16	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 17	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 18	LFOVS	LFOVS	PWJ	PWJ		

* Subject used semantically equivalent primitive operator

CPOTM	Compare Object to Memor
DS	Decision Step
LFOVS	Look for Object in Viewing Space
MHTKB	Move Hand to Keyboard
PB	Press Button
PWJ	Point with Joystick
RWS	Read What you See
TC	Type Code
TNL	Type Natural Language

Subsystem

- Cognitive Task
- Perceptual
- Physical
- Combined

Model 1	Method: Enter Quantity					
Baseline	MHTKB			TNL		PB
	Cog.	Motor	Cog.	Cog.	Motor	
Subject 1	MHTKB	MHTKB	TNL	TNL	TNL	
Subject 2			TNL	TNL	TNL	
Subject 3	MHTKB	MHTKB	TNL	TNL	TNL	
Subject 5	MHTKB	MHTKB	TNL	TNL	TNL	DS
Subject 6	MHTKB	MHTKB	TNL	TNL	TNL	
Subject 7			TNL	TNL	TNL	
Subject 8	MHTKB	MHTKB	TNL	TNL	TNL	
Subject 9			TNL*	TNL*	TNL*	DS
Subject 10	MHTKB	MHTKB	TNL	TNL	TNL	DS
Subject 11	MHTKB	MHTKB	TNL	TNL	TNL	DS
Subject 12	MHTKB	MHTKB	TNL	TNL	TNL	
Subject 13	MHTKB	MHTKB	TNL*	TNL*	TNL*	
Subject 14						
Subject 15	MHTKB	MHTKB	TNL*	TNL*	TNL*	
Subject 16	MHTKB	MHTKB	TNL	TNL	TNL	DS
Subject 17	MHTKB	MHTKB	TNL	TNL	TNL	DS
Subject 18			TNL	TNL	TNL	

* Subject used semantically equivalent primitive operator

CPOTM	Compare Object to Memory
DS	Decision Step
LFOVS	Look for Object in Viewing Space
MHTKB	Move Hand to Keyboard
PB	Press Button
PWJ	Point with Joystick
RWS	Read What you See
TC	Type Code
TNL	Type Natural Language

Subsystem
Cognitive
Perceptual
Physical
Combined

Model 1	Method: Verify			
Baseline	LFOVS		CPOTM	DS
	Perc.	Cog.	Cog.	Cog.
Subject 1				
Subject 2				
Subject 3				
Subject 5				
Subject 6				
Subject 7	LFOVS	LFOVS	CPOTM	DS
Subject 8	LFOVS*	LFOVS*	CPOTM	DS
Subject 9	LFOVS*	LFOVS*		
Subject 10	LFOVS*	LFOVS*	CPOTM	DS
Subject 11	LFOVS*	LFOVS*	CPOTM	DS
Subject 12				
Subject 13	LFOVS	LFOVS	CPOTM	DS
Subject 14	LFOVS	LFOVS	CPOTM	DS
Subject 15	LFOVS	LFOVS	CPOTM	DS
Subject 16	LFOVS	LFOVS	CPOTM	DS
Subject 17	LFOVS*	LFOVS*		
Subject 18				DS

* Subject used semantically equivalent primitive operator

CPOTM	Compare Object to Memory
DS	Decision Step
LFOVS	Look for Object in Viewing Space
MHTKB	Move Hand to Keyboard
PB	Press Button
PWJ	Point with Joystick
RWS	Read What you See
TC	Type Code
TNL	Type Natural Language

Cognitive
Perceptua
Physical
Combine

Summary of Accuracy Model 1 - Enemy Size								
	Cognitive		Perceptual		Physical		Total	
# of Steps	14		3		6		23	
	# of Steps Correct		# of Steps Correct		# of Steps Correct		# of Steps Correct	
Subject 1	9	64.29%	1	33.33%	6	100.00%	16	69.57%
Subject 2	8	57.14%	2	66.67%	4	66.67%	14	60.87%
Subject 3	8	57.14%	1	33.33%	6	100.00%	15	65.22%
Subject 5	10	71.43%	2	66.67%	6	100.00%	18	78.26%
Subject 6	9	64.29%	2	66.67%	6	100.00%	17	73.91%
Subject 7	11	78.57%	3	100.00%	5	83.33%	19	82.61%
Subject 8	13	92.86%	3	100.00%	6	100.00%	22	95.65%
Subject 9	10	71.43%	3	100.00%	5	83.33%	18	78.26%
Subject 10	13	92.86%	3	100.00%	6	100.00%	22	95.65%
Subject 11	14	100.00%	3	100.00%	6	100.00%	23	100.00%
Subject 12	5	35.71%	1	33.33%	3	50.00%	9	39.13%
Subject 13	9	64.29%	2	66.67%	3	50.00%	14	60.87%
Subject 14	10	71.43%	3	100.00%	4	66.67%	17	73.91%
Subject 15	13	92.86%	3	100.00%	6	100.00%	22	95.65%
Subject 16	14	100.00%	3	100.00%	6	100.00%	23	100.00%
Subject 17	12	85.71%	3	100.00%	6	100.00%	21	91.30%
Subject 18	8	57.14%	2	66.67%	4	66.67%	14	60.87%
Mean		73.95%		78.43%		86.27%		77.75%
Std. Dev.		18.02%		26.20%		18.85%		17.25%

Model 2	Method: Highlight Activity						
Baseline	LFOVS		DS	PWJ		PB	
	Perc.	Cog.	Cog.	Cog.	Motor	Cog.	Motor
Subject 1							
Subject 2							
Subject 3				PWJ	PWJ	PB	PB
Subject 4				PWJ	PWJ		
Subject 5	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 6	DS	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 7	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 8	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 9	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 10	LFOVS	LFOVS	DS	PWJ	PWJ	PB	PB
Subject 11	LFOVS*	LFOVS*	DS	PWJ	PWJ	PB	PB
Subject 12	LFOVS	LFOVS				PB	PB
Subject 13			DS				
Subject 14	DS	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 15	LFOVS	LFOVS	PWJ	PWJ	PB	PB	DS
Subject 16	LFOVS	LFOVS	DS	PWJ	PWJ	PB	PB
Subject 17	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 18	LFOVS	LFOVS		PWJ*	PWJ*	PB	PB

* Subject used semantically equivalent primitive operator

CPOTM	Compare Object to Memory
DS	Decision Step
LFOVS	Look for Object in Viewing Space
MHTKB	Move Hand to Keyboard
PB	Press Button
PWJ	Point with Joystick
RWS	Read What you See
TC	Type Code
TNL	Type Natural Language

Subsystem
- Cognitive
- Perceptual
- Physical
- Combined

Model 2	Method: Select Activity					
Baseline	LFOVS		PWJ		PB	
	Perc.	Cog.	Cog.	Motor	Cog.	Motor
Subject 1	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 2	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 3	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 4	LFOVS	LFOVS			PB	PB
Subject 5	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 6	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 7	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 8	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 9	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 10	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 11	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 12						
Subject 13	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 14	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 15	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 16	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 17	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 18	LFOVS	LFOVS	PWJ*	PWJ*	PB	PB

* Subject used semantically equivalent primitive operator

CPOTM	Compare Object to Memory
DS	Decision Step
LFOVS	Look for Object in Viewing Space
MHTKB	Move Hand to Keyboard
PB	Press Button
PWJ	Point with Joystick
RWS	Read What you See
TC	Type Code
TNL	Type Natural Languag

Subsystem

- Cognitive
- Perceptual
- Physical
- Combined

Model 2	Method: Verify			
Baseline	LFOVS		CPOTM	DS
	Perc.	Cog.	Cog.	Cog.
Subject 1				
Subject 2				
Subject 3				
Subject 4				
Subject 5				
Subject 6			CPOTM	DS
Subject 7				
Subject 8				
Subject 9	LFOVS*	LFOVS*	RWS	
Subject 10			RWS	
Subject 11	LAVS	LAVS	CPOTM	DS
Subject 12	LFOVS	LFOVS	CPOTM	
Subject 13				
Subject 14	LAVS	LAVS	CPOTM	DS
Subject 15	LFOVS*	LFOVS*	CPOTM	DS
Subject 16	LFOVS	LFOVS	CPOTM	DS
Subject 17	LFOVS*	LFOVS*		
Subject 18				

* Subject used semantically equivalent primitive operator

CPOTM	Compare Object to Memory
DS	Decision Step
LFOVS	Look for Object in Viewing Space
MHTKB	Move Hand to Keyboard
PB	Press Button
PWJ	Point with Joystick
RWS	Read What you See
TC	Type Code
TNL	Type Natural Language

Cognitive
Perceptua
Physical
Combine

Summary of Accuracy Model 2 - Enemy Activity								
	Cognitive		Perceptual		Physical		Total	
# of Steps	10		3		4		17	
	# of Steps Correct		# of Steps Correct		# of Steps Correct		# of Steps Correct	
Subject 1	3	30.00%	1	33.33%	2	50.00%	6	35.29%
Subject 2	3	30.00%	1	33.33%	2	50.00%	6	35.29%
Subject 3	5	50.00%	1	33.33%	4	100.00%	10	58.82%
Subject 4	3	30.00%	1	33.33%	2	50.00%	6	35.29%
Subject 5	6	60.00%	2	66.67%	4	100.00%	12	70.59%
Subject 6	9	90.00%	2	66.67%	4	100.00%	15	88.24%
Subject 7	6	60.00%	2	66.67%	4	100.00%	12	70.59%
Subject 8	6	60.00%	2	66.67%	4	100.00%	12	70.59%
Subject 9	8	80.00%	3	100.00%	4	100.00%	15	88.24%
Subject 10	8	80.00%	2	66.67%	4	100.00%	14	82.35%
Subject 11	10	100.00%	3	100.00%	4	100.00%	17	100.00%
Subject 12	4	40.00%	2	66.67%	1	25.00%	7	41.18%
Subject 13	4	40.00%	1	33.33%	2	50.00%	7	41.18%
Subject 14	10	100.00%	3	100.00%	4	100.00%	17	100.00%
Subject 15	10	100.00%	3	100.00%	4	100.00%	17	100.00%
Subject 16	10	100.00%	3	100.00%	4	100.00%	17	100.00%
Subject 17	7	70.00%	3	100.00%	4	100.00%	14	82.35%
Subject 18	6	60.00%	2	66.67%	4	100.00%	12	70.59%
Mean		65.56%		68.52%		84.72%		70.59%
Std. Dev.		0.2572		0.2675		0.2592		0.2421

Model 3	Method: Highlight Location						
Baseline	LFOVS		DS	PWJ		PB	
	Perc.	Cog.	Cog.	Cog.	Motor	Cog.	Motor
Subject 1				PWJ	PWJ	PB	PB
Subject 2	LFOVS	LFOVS		PWJ*	PWJ*		
Subject 3	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 4	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 5	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 6	DS	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 7	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 8	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 9	LFOVS	LFOVS	DS	PWJ	PWJ	PB	PB
Subject 10	LFOVS	LFOVS	DS	PWJ	PWJ	PB	PB
Subject 11	LAVS	LAVS	DS	PWJ	PWJ	PB	PB
Subject 12	LFOVS	LFOVS				PB	PB
Subject 13			DS				
Subject 14	DS	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 15	LFOVS	LFOVS	PWJ	PWJ	PB	PB	DS
Subject 16	LFOVS	LFOVS	DS	PWJ	PWJ	PB	PB
Subject 17	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 18							

* Subject used semantically equivalent primitive operator

CPOTM	Compare Object to Memory
DS	Decision Step
LFOVS	Look for Object in Viewing Space
MHTKB	Move Hand to Keyboard
PB	Press Button
PWJ	Point with Joystick
RWS	Read What you See
TC	Type Code
TNL	Type Natural Language

Subsystem
- Cognitive
- Perceptual
- Physical
- Combined

Model 3	Method: Enter Enemy Location					
Baseline	LFOVS		PWJ		PB	
	Perc.	Cog.	Cog.	Motor	Cog.	Motor
Subject 1			PB	PB	PWJ	PWJ
Subject 2	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 3	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 4	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 5	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 6	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 7	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 8	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 9	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 10	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 11	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 12	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 13	LFOVS*	LFOVS*	PWJ	PWJ	PB	PB
Subject 14	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 15	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 16	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 17	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 18	LFOVS	LFOVS	PWJ	PWJ	PB	PB

* Subject used semantically equivalent primitive operator

CPOTM	Compare Object to Memory
DS	Decision Step
LFOVS	Look for Object in Viewing Space
MHTKB	Move Hand to Keyboard
PB	Press Button
PWJ	Point with Joystick
RWS	Read What you See
TC	Type Code
TNL	Type Natural Languag

Subsystem
- Cognitive
- Perceptual
- Physical
- Combined

Model 3	Method: Enter Observer's Location					
Baseline	LFOVS		PWJ		PB	
	Perc.	Cog.	Cog.	Motor	Cog.	Motor
Subject 1						
Subject 2	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 3	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 4	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 5	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 6	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 7	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 8	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 9	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 10	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 11	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 12	LFOVS	LFOVS	PWJ	PWJ		
Subject 13			PWJ	PWJ	PB	PB
Subject 14	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 15	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 16	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 17	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 18						

* Subject used semantically equivalent primitive operator

CPOTM	Compare Object to Memory
DS	Decision Step
LFOVS	Look for Object in Viewing Space
MHTKB	Move Hand to Keyboard
PB	Press Button
PWJ	Point with Joystick
RWS	Read What you See
TC	Type Code
TNL	Type Natural Languag

Subsystem
- Cognitive
- Perceptual
- Physical
- Combined

Model 3	Method: Enter Course					
Baseline	LFOVS		PWJ		PB	
	Perc.	Cog.	Cog.	Motor	Cog.	Motor
Subject 1						
Subject 2	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 3	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 4	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 5	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 6	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 7	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 8	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 9	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 10	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 11	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 12	LFOVS	LFOVS	PWJ	PWJ		
Subject 13	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 14	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 15	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 16	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 17	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 18						

* Subject used semantically equivalent primitive operator

CPOTM	Compare Object to Memory
DS	Decision Step
LFOVS	Look for Object in Viewing Space
MHTKB	Move Hand to Keyboard
PB	Press Button
PWJ	Point with Joystick
RWS	Read What you See
TC	Type Code
TNL	Type Natural Languag

Subsystem

- Cognitive
- Perceptual
- Physical
- Combined

Model 3	Method: Enter Speed										
	LFOVS		PWJ		PB		MHTKB		TNL		
	P	C	C	M	C	M	C	P	C	C	M
Subject 1											
Subject 2	LFOVS	LFOVS	PWJ	PWJ					TNL	TNL	TNL
Subject 3	LFOVS	LFOVS	PWJ	PWJ	PB	PB	MHTKB	MHTKB	TNL	TNL	TNL
Subject 4	LFOVS	LFOVS	PWJ	PWJ	PB	PB					
Subject 5	LFOVS	LFOVS	PWJ	PWJ	PB	PB	MHTKB	MHTKB	TNL	TNL	TNL
Subject 6	LFOVS	LFOVS	PWJ	PWJ	PB	PB	MHTKB	MHTKB	TNL	TNL	TNL
Subject 7	LFOVS	LFOVS	PWJ	PWJ	PB	PB					
Subject 8	LFOVS	LFOVS	PWJ	PWJ	PB	PB	MHTKB	MHTKB	TNL	TNL	TNL
Subject 9	LFOVS	LFOVS					MHTKB	MHTKB	TNL*	TNL	TNL
Subject 10	LFOVS	LFOVS	PWJ	PWJ	PB	PB	MHTKB	MHTKB	TNL	TNL	TNL
Subject 11	LFOVS	LFOVS	PWJ	PWJ	PB	PB	MHTKB	MHTKB	TNL	TNL	TNL
Subject 12	LFOVS	LFOVS	PWJ	PWJ			MHTKB	MHTKB			
Subject 13	LFOVS	LFOVS	PWJ	PWJ					TNL*	TNL	TNL
Subject 14	LFOVS	LFOVS	PWJ	PWJ	PB	PB	MHTKB	MHTKB	TNL	TNL	TNL
Subject 15	LFOVS	LFOVS	PWJ	PWJ	PB	PB			TNL*	TNL	TNL
Subject 16	LFOVS	LFOVS	PWJ	PWJ	PB	PB	MHTKB	MHTKB	TNL	TNL	TNL
Subject 17	LFOVS	LFOVS	PWJ	PWJ	PB	PB	MHTKB	MHTKB	TNL	TNL	TNL
Subject 18	LFOVS	LFOVS	PWJ*	PWJ*					TNL*	TNL	TNL

* Subject used semantically equivalent primitive operator

CPOTM	Compare Object to Memory
DS	Decision Step
LFOVS	Look for Object in Viewing Space
MHTKB	Move Hand to Keyboard
PB	Press Button
PWJ	Point with Joystick
RWS	Read What you See
TC	Type Code
TNL	Type Natural Language

Subsystem
- Cognitive
- Perceptual
- Physical
- Combined

Model 3	Method: Verify			
Baseline	LFOVS		CPOTM	DS
	Perc.	Cog.	Cog.	Cog.
Subject 1				
Subject 2				
Subject 3				
Subject 4				
Subject 5	LFOVS	LFOVS		
Subject 6			CPOTM	DS
Subject 7				
Subject 8	LAVS	LAVS	CPOTM	DS
Subject 9	LFOVS	LFOVS	CPOTM*	
Subject 10	LAVS	LAVS	CPOTM	DS
Subject 11	LAVS	LAVS	CPOTM	DS
Subject 12	LAVS	LAVS	CPOTM	DS
Subject 13			CPOTM	
Subject 14	LAOVS	LAOVS	CPOTM	DS
Subject 15	LFOVS	LFOVS	CPOTM	DS
Subject 16	LFOVS	LFOVS	CPOTM	DS
Subject 17	LFOVS*	LFOVS*		
Subject 18	LFOVS	LFOVS	CPOTM	

* Subject used semantically equivalent primitive operator

CPOTM	Compare Object to Memory
DS	Decision Step
LFOVS	Look for Object in Viewing Space
MHTKB	Move Hand to Keyboard
PB	Press Button
PWJ	Point with Joystick
RWS	Read What you See
TC	Type Code
TNL	Type Natural Language

Cognitive
Perceptua
Physical
Combine

Summary of Accuracy Model 3 - Location								
	Cognitive		Perceptual		Physical		Total	
# of Steps	22		6		12		40	
	# of Steps Correct		# of Steps Correct		# of Steps Correct		# of Steps Correct	
Subject 1	4	18.18%	0	0.00%	4	33.33%	8	20.00%
Subject 2	15	68.18%	5	83.33%	9	75.00%	29	72.50%
Subject 3	18	81.82%	5	83.33%	12	100.00%	35	87.50%
Subject 4	15	68.18%	5	83.33%	10	83.33%	30	75.00%
Subject 5	19	86.36%	6	100.00%	12	100.00%	37	92.50%
Subject 6	21	95.45%	5	83.33%	12	100.00%	38	95.00%
Subject 7	15	68.18%	5	83.33%	10	83.33%	30	75.00%
Subject 8	21	95.45%	6	100.00%	12	100.00%	39	97.50%
Subject 9	19	86.36%	6	100.00%	10	83.33%	35	87.50%
Subject 10	22	100.00%	6	100.00%	12	100.00%	40	100.00%
Subject 11	22	100.00%	6	100.00%	12	100.00%	40	100.00%
Subject 12	15	68.18%	6	100.00%	7	58.33%	28	70.00%
Subject 13	14	63.64%	3	50.00%	8	66.67%	25	62.50%
Subject 14	22	100.00%	6	100.00%	12	100.00%	40	100.00%
Subject 15	21	95.45%	6	100.00%	11	91.67%	38	95.00%
Subject 16	22	100.00%	6	100.00%	12	100.00%	40	100.00%
Subject 17	19	86.36%	6	100.00%	12	100.00%	37	92.50%
Subject 18	9	40.91%	3	50.00%	4	33.33%	16	40.00%
Mean		79.04%		84.26%		83.80%		81.25%
Std. Dev.		0.2249		0.2649		0.2241		0.2232

Model 4	Method: Highlight Unit						
Baseline	LFOVS		DS	PWJ		PB	
	Perc.	Cog.	Cog.	Cog.	Motor	Cog.	Motor
Subject 1	PWJ	PWJ		LFOVS	LFOVS	PB	PB
Subject 2	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 3	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 4							
Subject 5	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 6	DS	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 7	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 8	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 9	LFOVS	LFOVS	DS	PWJ	PWJ	PB	PB
Subject 10	LFOVS	LFOVS	DS	PWJ	PWJ	PB	PB
Subject 11	LFOVS	LFOVS	DS	PWJ	PWJ	PB	PB
Subject 12	LFOVS	LFOVS		PWJ	PWJ		
Subject 13	LFOVS	LFOVS	DS	PWJ	PWJ	PB	PB
Subject 14	DS	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 15	LFOVS	LFOVS	PWJ	PWJ	PB	PB	DS
Subject 16	LFOVS	LFOVS	DS	PWJ	PWJ	PB	PB
Subject 17	LFOVS	LFOVS		PWJ	PWJ	PB	PB
Subject 18	LFOVS	LFOVS		PWJ	PWJ	PB	PB

* Subject used semantically equivalent primitive operator

CPOTM	Compare Object to Memory
DS	Decision Step
LFOVS	Look for Object in Viewing Space
MHTKB	Move Hand to Keyboard
PB	Press Button
PWJ	Point with Joystick
RWS	Read What you See
TC	Type Code
TNL	Type Natural Language

Subsystem
Cognitive
Perceptual
Physical
Combined

Model 3	Method: Enter Unit										
	LFOVS		PWJ		PB		MHTKB		TNL		
	P	C	C	M	C	M	C	P	C	C	M
Subject 1	LFOVS	LFOVS			PB	PB	MHTKB	MHTKB	TNL	TNL	TNL
Subject 2	LFOVS	LFOVS	PWJ	PWJ	PB	PB			TNL	TNL	TNL
Subject 3	LFOVS	LFOVS	PWJ	PWJ	PB	PB	MHTKB	MHTKB	TNL	TNL	TNL
Subject 4	LFOVS	LFOVS	PWJ	PWJ	PB	PB					
Subject 5	LFOVS	LFOVS	PWJ	PWJ	PB	PB	MHTKB	MHTKB	TNL	TNL	TNL
Subject 6	LFOVS	LFOVS	PWJ	PWJ	PB	PB	MHTKB	MHTKB	TNL	TNL	TNL
Subject 7	LFOVS	LFOVS	PWJ	PWJ	PB	PB			TNL	TNL	TNL
Subject 8	LFOVS	LFOVS	PWJ	PWJ	PB	PB	MHTKB	MHTKB	TNL	TNL	TNL
Subject 9	LFOVS	LFOVS	PWJ	PWJ	PB	PB	MHTKB	MHTKB	TNL*	TNL	TNL
Subject 10	LFOVS	LFOVS	PWJ	PWJ	PB	PB	MHTKB	MHTKB	TNL	TNL	TNL
Subject 11	LFOVS	LFOVS	PWJ	PWJ	PB	PB	MHTKB	MHTKB	TNL	TNL	TNL
Subject 12	LFOVS	LFOVS	PWJ	PWJ			MHTKB	MHTKB			
Subject 13											
Subject 14											
Subject 15	LFOVS	LFOVS	PWJ	PWJ	PB	PB			TNL*	TNL	TNL
Subject 16	LFOVS	LFOVS	PWJ	PWJ	PB	PB	MHTKB	MHTKB	TNL	TNL	TNL
Subject 17	LFOVS	LFOVS	PWJ	PWJ	PB	PB	MHTKB	MHTKB	TNL	TNL	TNL
Subject 18	LFOVS	LFOVS	PWJ	PWJ	PB	PB					

* Subject used semantically equivalent primitive operato

CPOTM	Compare Object to Memory
DS	Decision Step
LFOVS	Look for Object in Viewing Space
MHTKB	Move Hand to Keyboard
PB	Press Button
PWJ	Point with Joystick
RWS	Read What you See
TC	Type Code
TNL	Type Natural Language

Subsystem

Cognitive

Perceptual

Physical

Combined

Model 4	Method: Highlight Unit Type					
Baseline	LFOVS		PWJ		PB	
	Perc.	Cog.	Cog.	Motor	Cog.	Motor
Subject 1	LFOVS	LFOVS			PB	PB
Subject 2	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 3	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 4						
Subject 5	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 6	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 7	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 8	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 9	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 10	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 11	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 12	LFOVS	LFOVS	PWJ	PWJ		
Subject 13						
Subject 14	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 15	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 16	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 17	LFOVS	LFOVS	PWJ	PWJ	PB	PB
Subject 18	LFOVS	LFOVS	PWJ	PWJ	PB	PB

* Subject used semantically equivalent primitive operator

CPOTM	Compare Object to Memory
DS	Decision Step
LFOVS	Look for Object in Viewing Space
MHTKB	Move Hand to Keyboard
PB	Press Button
PWJ	Point with Joystick
RWS	Read What you See
TC	Type Code
TNL	Type Natural Languag

Subsystem

- Cognitive
- Perceptual
- Physical
- Combined

Model 4	Method: Verify			
Baseline	LFOVS		CPOTM	DS
	Perc.	Cog.	Cog.	Cog.
Subject 1				
Subject 2				
Subject 3				
Subject 4				
Subject 5				
Subject 6				DS
Subject 7	LFOVS	LFOVS	CPOTM	DS
Subject 8	LAVS	LAVS	CPOTM	DS
Subject 9	LFOVS*	LFOVS*	RWS	
Subject 10	LAVS	LAVS	CPOTM	DS
Subject 11	LAVS	LAVS	CPOTM	DS
Subject 12	LFOVS	LFOVS		DS
Subject 13			CPOTM	
Subject 14	LAVS	LAVS	CPOTM	DS
Subject 15	LFOVS	LFOVS	CPOTM	DS
Subject 16	LFOVS*	LFOVS*	CPOTM	DS
Subject 17	LFOVS*	LFOVS*		
Subject 18	LFOVS	LFOVS		

* Subject used semantically equivalent primitive operator

CPOTM	Compare Object to Memory
DS	Decision Step
LFOVS	Look for Object in Viewing Space
MHTKB	Move Hand to Keyboard
PB	Press Button
PWJ	Point with Joystick
RWS	Read What you See
TC	Type Code
TNL	Type Natural Language

Cognitive
Perceptua
Physical
Combine

Summary of Accuracy Model 4 – Enemy Unit								
	Cognitive		Perceptual		Physical		Total	
# of Steps	16		4		8		28	
	# of Steps Correct		# of Steps Correct		# of Steps Correct		# of Steps Correct	
Subject 1	10	62.50%	3	75.00%	6	75.00%	19	67.86%
Subject 2	11	68.75%	3	75.00%	7	87.50%	21	75.00%
Subject 3	12	75.00%	3	75.00%	8	100.00%	23	82.14%
Subject 4	3	18.75%	1	25.00%	2	25.00%	6	21.43%
Subject 5	12	75.00%	3	75.00%	8	100.00%	23	82.14%
Subject 6	14	87.50%	3	75.00%	8	100.00%	25	89.29%
Subject 7	14	87.50%	4	100.00%	7	87.50%	25	89.29%
Subject 8	15	93.75%	4	100.00%	8	100.00%	27	96.43%
Subject 9	15	93.75%	4	100.00%	8	100.00%	27	96.43%
Subject 10	16	100.00%	4	100.00%	8	100.00%	28	100.00%
Subject 11	16	100.00%	4	100.00%	8	100.00%	28	100.00%
Subject 12	9	56.25%	4	100.00%	4	50.00%	17	60.71%
Subject 13	5	31.25%	1	25.00%	2	25.00%	8	28.57%
Subject 14	10	62.50%	3	75.00%	4	50.00%	17	60.71%
Subject 15	15	93.75%	4	100.00%	7	87.50%	26	92.86%
Subject 16	16	100.00%	4	100.00%	8	100.00%	28	100.00%
Subject 17	13	81.25%	4	100.00%	8	100.00%	25	89.29%
Subject 18	10	62.50%	4	100.00%	6	0.75	20	0.714285714
Mean		75.00%		83.33%		81.25%		77.98%
Std. Dev.		0.2329		0.2425		0.2617		0.2322

REFERENCES

Army Digitization Office. (1999). *Defining Digitization.* [On-line Presentation] (Available at www.ado.army.mil)

Brusilovsky, P.L. (1994). The Construction and Application of Student Models in Intelligent Tutoring Systems. *Journal of Computer and Systems Sciences International.* Silver Springs: Scripta Technica, Inc., 70-89)

Card, Moran, & Newell. (1983) *The Psychology of Human-Computer Interaction.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Cohen, Jacob and Cohen Patricia (1975). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Dillenbourg, P. and J.A. Self. (1992). PEOPLE POWER: A human-computer collaborative learning system. *Proceeding of ITS*, held in Montreal, 651-660.

Finney, D.J. (1960). *An Introduction to the Theory of Experimental Design.* The University of Chicago Press.

Kieras, D.E. (1988). Towards a practical GOMS model methodology for user interface design. In M. Helander (Ed.), *Handbook of Human-Computer Interaction* , 135-157.

Kieras, D.E., & Polson, P.G. (1985). An approach to the formal analysis of user complexity. *International Journal of Man-Machine Studies, 22*, 365-394

King, R.L., Classification of student modeling approaches for intelligent tutoring. *MSSU-COE-ERC-98-4*, Mississippi State University, Jan. 1998.

Kotnour, Timothy G., (1992). *Design, Development, and Testing of an Automated Knowledge Acquisition Tool to Aid Problem Solving, Decision Making, and Planning.* Unpublished master's thesis, Virginia Polytechnic Institute and State University.

Mendenhall, William, and Sincich Terry (1995). Statistics for Engineering and the Sciences. Upper Saddle River, New Jersey: Prentice Hall

Myer, D.E. & Kieras, D. E. (1994). *EPIC computational models of psychological refractory-period effects in human multiple-task performance.* Office of Naval Research Technical Report, Contract Number N00014-92-J-1173

Sanders, William R., (1999). *Technical Report 1096: Digital Procedural Skill Retention for Selected M1A2 Tank Inter-Vehicular Information System (IVIS) Tasks.* United States Army Research Institute for the Behavioral and Social Sciences.

Siegel, Sidney, (1956). *Nonparametric Statistics for the Behavioral Sciences.* McGraw-Hill Book Company.

Steinberg, Esther R. (1984). *Teaching Computers to Teach. Computer-based education.* Research Laboratory and College of Education University of Illinois at Urbana-Champaign. Hillsdale, NJ: Lawrence Earlbaum Associates, Publishers.

United Defense, Limited Partners. (15 March 1999). *BDT Student Guide,* v3.0z SW Release

United States Army Infantry Center, Directorate of Training. (1999). *System Training Plan (STRAP)M2A3/M3A3 Bradley Fighting Vehicle.* Fort Benning, GA (1999)

United States Army Infantry Center, Directorate of Training. (1999). *Bradley Fighting Vehicle System Modernization Operational Requirements Document, Appendix B – Training Support Requirements.* Fort Benning, GA

Williams, Kent E. (1998). *An Automated Aid for the Conduct of a Detailed Cognitive Task Analysis for Modeling Human Computer Interaction Performance.* The Office of Naval Research Cognative Science Program.

Williams, Kent E., and Kotnor, Tim G. (1993). *Knowledge Acquisition: A Review of Manual, Machine-Aided and Machine Learning Methods.* Virginia Polytechnic Institute and State University. Interim Technical Report.

Williges, B.H. & Williges, R.C. (1984). Dialogue design for consideration for interactive computer systems. *Human Factors Review:* 1984, 167-208